2019-09

Working paper. Economics

ISSN 2340-5031

INVESTMENT CLIMATE EFFECTS ON ALTERNATIVE FIRM-LEVEL PRODUCTIVITY MEASURES

ALVARO ESCRIBANO, J. LUIS GUASCH AND JORGE PENA

uc3m | Universidad Carlos III de Madrid

INVESTMENT CLIMATE EFFECTS ON ALTERNATIVE FIRM-LEVEL PRODUCTIVITY MEASURES*

ALVARO ESCRIBANO,† J. LUIS GUASCH AND JORGE PENA‡

February , 2019

Abstract Developing countries are increasingly concerned about improving country competitiveness and productivity. Investment Climate surveys (ICs) at the firm level, are becoming the standard way for the World Bank to identify key obstacles to country competitiveness. This paper develops a general to specific econometric methodology, based on firm level observable fixed effects that generate robust investment climate effects (elasticities) on total factor productivity (TFP). By robust IC elasticities on TFP we mean elasticity estimates with equal signs and of similar magnitudes for several competing TFP measures. We apply this econometric methodology to the IC survey of Costa Rica showing how robust the investment climate effects are for several measures of TFP when conditioning on relevant plant-level information that is usually unobserved. For the economic evaluation we estimate the marginal effects of each IC variable on TFP as well as their IC impacts on average TFP obtaining important economic differences. These IC estimates are obtained from five blocks of IC variables, (i) infrastructure, (ii) red tape, corruption and crime, (iii) finance and corporate governance, (iv) quality, innovation and labor skills and (v) other control variables, could be used as benchmarks to assess cross-country IC assessments of TFP.

Key words: Total factor productivity, investment climate, observable fixed effects, robust estimates, input-output elasticities, impact evaluation on average TFP, demeaned TFP.

JEL Codes: C23, C18, L25, L11, F14, C51

---

* This paper is based on unpublished work done by Escribano, Guasch and Pena on the investment climate assessment (ICA) of the World Bank for more than 48 developing countries. We are indebted to D. Ackerberg, J.M. Dufour, A. Pakes , J. Levinsohn for the suggestions given on previous versions of this paper. We have also benefited from the suggestions of participants of the American Association meetings, from seminars and courses given by A. Escribano at the World Bank, at CORE (UCL, Belgium) and from the participants of the courses given at the University Carlos III de Madrid. Alvaro Escribano and Jorge Pena acknowledges funding from The World Bank. Alvaro Escribano acknowledge funding from the Spanish Ministry of Economy, Industry and Competitiveness (ECO2015-68715-R, ECO2016-00105-001), Consolidation Grant (#2006/04046/002), and Maria de Maeztu Grant (MDM 2014-0431).
† Department of Economics, Universidad Carlos III de Madrid Chair on Internationalization. Corresponding autor: Alvaro Escribano, Department of Economics, Universidad Carlos III de Madrid, Calle Madrid 126, 28903 Madrid, Spain. Phone 34-91-6249879, Fax 34-91-6248908. alvaroe@eco.uc3m.es.
‡ IE University, Madrid, Spain.

1 Introduction

As countries face the pressures and impacts of globalization, they are seeking ways to stimulate growth and employment within this context of increased openness. Developing countries are focusing on issues of competitiveness and total factor productivity (TFP) through microeconomic reform programs. From South East Asia to Latin America, countries are reformulating their strategies and making increased competitiveness a key priority of government programs.

Prescott (1998) argues that to understand large international income differences, it is necessary to explain differences in productivity (TFP). His main candidates to explain those gaps is the resistance to the adoption of new technologies and to the efficient use of current operating technologies, which in turn are conditioned by the institutional and policy arrangements a society employs (the investment climate for us). Cole et al. (2004) also have argued that Latin America has not replicated Western economic success due to the productivity (TFP) gap. They point to competitive barriers as the promising channels for understanding the low productivity observed in Latin American countries. It is now well accepted conceptually and empirically, that the scope and nature of regulations on economic activity and factor markets—the so-called investment climate and business environment—can significantly and adversely impact productivity, growth and economic activity.[4] A significant component of country competitiveness is having a *"good investment climate" or "business environment"*.

The investment climate (IC), as defined in the World Development Report, World Bank (2005), is "*the set of location-specific factors shaping the opportunities and incentives for firms to invest productively, create jobs and expand*". In this paper we are able to measure the IC effects on TFP at the plant-level, using data from investment climate surveys (ICs) of the World Bank. These surveys are stratified random samples of firms, mainly *manufacturing firms*, with stratification variables being industry, region and size. The sampling processes are done in close partnership with regional statistical agencies that provided the necessary information on the total census of manufacturing firms in

[4] See Bosworth and Collins, 2003; Rodrik and Subramanian, 2004; McMillan, 1998 and 2004; OECD, 2001; Wilkinson, 2001; Alexander et al., 2004; Djankov et al., 2002; Haltiwanger, 2002; He et al., 2003; Dethier et al (2008) and World Bank, 2003, 2004a,b and 2005.

each country,[5] keeping the basic structure and questions of each investment climate survey common to all the countries. These ICs represent a very rich quantitative source of plant-level information, *usually unobservable fixed effects*, that would allow us to study new determinants of business environments and new sources of bottlenecks for firm´s growth.

The ICs for each developing country is a unbalanced panel of a large number of firms with two important characteristics: a) few years of temporal (three years of recall data) observations on plant level production function variables and b) only one year of plant level information on their investment climate (see Tables B.1 to B.2 of Appendix B for the list and definitions of all significant IC variables of Costa Rica). We assume that, unless important structural breaks occurs in the economy, these IC values at the plant-level should not change much from the two or three previous consecutive years. For example, consider the infrastructure IC variable named *number of power outages* suffered by a firm in a given year. Since the quality of the electricity system is given in the short run, the firm´s expected number of power outages should be almost constant (fixed effect) during few consecutive years. Therefore in our IC data base, for each i-plant we assume that the reported IC_i values for the last year are preserved for the previous two years creating a plant level *matrix of observable fixed effects*[6].

Hall and Jones (1999) argue that to explain differences in levels of long-run economic success across countries, one is forced to focus on more basic determinants: like infrastructure and persistent barriers that make technology and capital not moving fast across borders, and continue saying that *"... long-run determinants of economic success are factors that are changing slowly over time"*. For us, those determinants are associated here with plant-level observable fixed effects related to the investment climate (IC). In fact, we have a rich plant-level information on infrastructure, red tape, corruption, finance, innovation, labor skills, competitive environment, etc., see Tables B.1 and B.2 of Appendix B for more information.

---

[5] In order to ensure enough number of large establishments in the sample of manufacturing firms, a sampling approach which oversample large firms was applied.

[6] We were suggested by J. Levinsohn to compare our results with those obtained using only cross section data, without repeating the values of the IC values during few years. Almost identical but less efficient parameter estimates were obtained when using only cross section data instead of recall data (results are available upon request).

The TFP methodology of this paper was developed to explain why different researchers[7] addressing common issues related to the IC effects of infrastructure and finance on TFP, were reaching opposite conclusions (different signs in the coefficients of key IC variables and selecting different IC variables) using the same data from IC surveys. Four main possible different methodological sources were under consideration as the main causes for getting opposite IC signs on TFP: 1) the decision to pool or not to pool surveys data from different countries, 2) the different level of aggregation considered within a country (industry or country level coefficients), 3) the different TFP estimation procedures considered (GMM, 2SLS, 2-step Solow´s residuals, structural approaches, like Olley and Pakes´s approach, etc.), 4) the different functional forms of the production functions considered (Translog, Cobb-Douglas, etc.).

The main questions address in this paper are the following; do the signs and order of magnitude of the IC elasticity estimates on TFP crucially depended on the particular TFP measure used or on the alternative TFP estimation procedure considered? Which econometric methodology could different researchers use to evaluate the IC impact on TFP, avoiding getting contradictory conclusions in terms of size and orders of magnitude of the IC coefficients on TFP using the same data set?

We show that to get the correct signs of the IC coefficients on TFP and similar s order of magnitude, almost any of the reasonable alternative measures of TFP could be used, as long as the researcher controls for the relevant plant-level IC information affecting firm´s decisions. Those key unobserved variables are usually treated as fixed or random effects by the econometricians. In this paper, we show that IC surveys provide a very rich firm level information, usually associated with unobserved effects, that allow us to get IC elasticities on TFP that are robust (equal signs and similar order of magnitude) for alternative TFP measures and alternative estimation procedures, including ordinary least squares (OLS). We show that the orders of magnitude of the IC coefficients on TFP are similar. Particular questions related to IC differences in magnitude on TFP are standard and easily testable when we are looking for the true (or best) model.

[7] In fact, the different researchers were part of different units (infrastructure, finance, etc.) of the World Bank (WB) in Washington DC. To answer these methodological questions the WB lunched an open-call for econometric proposals. Several proposal were suggested. The econometric methodology of this paper was selected since it was the only one able to encompass previous contradictory results and identify the cause for getting opposite empirical results.

To get robust IC signs on TFP, we suggest that a simple extended production function approach could be used in which the firm specific productivity shocks, usually observed by the managers but not by the econometrician, are initially proxied by more than 150 plant-level investment climate variables. These initially large number of explanatory IC variables are later on reduced to 26 significant IC variables, after using a combination of general to specific approach and specific to general testing procedure at the end to make sure we are not omitting any relevant IC variables.

When any of the production function inputs is influenced by common causes affecting TFP, like IC variables or other plant characteristics, there is a simultaneous equation problem, resulting in the well-known transmission bias of OLS estimators in production functions.[8] We overcome this simultaneous equation problem we estimate the IC effects on the TFP using a panel of manufacturing firms of Costa Rica for years 2002, 2003 and 2004 but controlling for plant-level IC observable fixed-effects.

The development of this *robust TFP specification strategy* that could be used as a benchmark for comparison of alternative studies of the impact of IC variables on firm´s productivity, is another objective of this paper. By robust TFP methodology we mean one that provides similar IC results on TFP; with equal signs and of similar order of magnitude for several competing TFP measures. This property is essential to make cross-country comparisons.[9]

In particular, our estimates of Costa Rica are robust across eight different TFP measures coming from: 1) *different functional forms* of the production functions, 2) *different TFP estimates* and 3) *different levels of aggregation* of the *input-output elasticities* (at industry

---

[8] There is an extensive literature discussing the advantages and disadvantages of using different statistical estimation techniques and/or growth accounting (index number) techniques to estimate productivity or Total Factor Productivity in levels (TFP) or in rates of growth (TFPG). For overviews of different productivity concepts and aggregation alternatives see Solow (1957), Jorgenson, Gollop and Fraumeni (1987), Hall (1990), Olley and Pakes (1996), Foster, Haltiwanger and Krizan (1998), Batelsman and Doms (2000), Hulten (2001), Diewert and Nakamura (2002), Jorgenson (2001), Barro and Sala-i-Martin (2004) and Ackerberg, Benkard, Berry and Pakes, (2007).

[9] This methodology has been applied in background documents on investment climate assessment *(ICA) of the World Bank* covering 42 developing countries. This list of countries includes Eritrea, Ethiopia, Madagascar, Malawi, Niger, Tanzania, Zambia, Burkina Faso, Uganda, Mali, Kenya, Senegal, Mauritania, Bangladesh, Honduras, Pakistan, Cameroon, India, Bolivia, Guatemala, Honduras, Nicaragua, Philippines, Morocco, Indonesia, Ecuador, El Salvador, Egypt, Namibia, Turkey, Algeria, Colombia, Brazil, Mexico, Botswana, Costa Rica, South Africa, Swaziland, Croatia, Chile, Mauritius, Pakistan and Peru. The robustness of the TFP results that we present here for Costa Rica as an example are maintained in all these countries.

and country level).[10]Three approaches were considered to show evidence in favor of the *robustness* hypothesis of the IC empirical results on TFP. *First,* showing that the signs of all the IC effects on TFP are always the same for all TFP measures. The results for the 26 coefficients on IC variables show that the signs of the "IC elasticities on TFP" are all the same (except for the unrestricted Translog where only one IC coefficient change the sign[11]). *Second,* testing that the magnitudes of all of the 26 IC coefficients on TFP are equal to those of our benchmark model for most TFP measures. The results show that in four out of the seven TFP measures, the magnitudes of the IC coefficients do not vary for more than 92% of them. However, using Levinshon and Petrin (2003) (LP) and Ackerberg, Caves and Frazer (2006) (AC&F) structural estimators this percentage is reduced to 54%. The reason is that impacts of certain IC variables compete with the extra nonlinear input-terms added in these procedures as proxies for unobserved firm´s productivity. *Third,* testing that the densities and cumulative distributions of TFP, after being demeaned, are equal for all TFP measures. Out of the eight TFP measures considered the main significant differences only show up in the unrestricted (by industry) cases and in the Translog.

A summary of the main empirical results obtained for the TFP analysis of the investment climate (IC) of Costa Rica is the following: a) IC represents 72.3% of the contribution to average TFP. Therefore improving the IC is important to enhance TFP in Costa Rica. b) The ranking of the IC´s contributions to average TFP by blocks of IC variables is the following: 1) "Red Tape, Corruption and Crime" with the highest contribution of 34.4% of the total, 2) "Finance and Corporate Governance" with 22.9% contribution, 3) "Infrastructures" with 17.9%, 4) "Other Control variables" with 17,4% and finally 5) " Quality, Innovation and Labor Skills" with a 7.3%. c) The single most important IC variable in terms of TFP in Costa Rica is related to *informality* of the firms. In particular, the elasticity of "Sales Declared for Taxes" on TFP is positive and small (0.01) but its contribution to average TFP is the largest (19.6%). The reason is that informality affects many of the firms of Costa Rica´s manufacturing sector affecting therefore the average

[11] The unrestricted (by industry) Translog production function has too many parameters and it is known to provide unstable numerical results (due to multicolinearity, etc.). However, we show that by using our econometric methodology only one of the signs of only one of the IC coefficients changes, but this variable was not significant.

TFP level of the country. The second most important IC constraint is “Wait for Electricity Supply” with a negative elasticity equal to -0.128 but with a 9.7% contribution to average TFP.

The paper is structured as follows. Section 2 introduces the concepts of productivity (TFP) and discusses general productivity measures based on levels versus differences. We conclude that, given the fixed effect nature of IC variables obtained form ICs, it is better to analyze productivity in levels (or log-levels) rather than rates of growth of productivity. A specific solution to the endogeneity problem of the inputs of the production function will be presented in section 2.2 when estimation issues of production functions are discussed. This section also introduces an econometric methodology for a robust selection of IC and firm explanatory variables for different productivity (TFP) measures. Section 3, presents seven additional econometric specifications along with a statistical test of equality of IC parameters across specifications to check the robustness of the results of the baseline or benchmark model proposed in section 2.2. In section 4 we present the main empirical results regarding the investment climate conditions obtained in the paper. This section also suggests evaluating the country specific contribution of IC variables on average productivity. Finally, section 5 presents a summary of the econometric methodology and of the main conclusions. All the Figures and Tables with the definitions of the variables used and with the panel data estimation results are included in the Appendix.

2 TFP Measures when Investment Climate variables are observable fixed effects

Since there is no single salient measure of TFP_{it}, therefore any empirical evaluation on the productivity (TFP) impact of IC variables might critically depend on the way productivity is measured. To avoid having to select a particular TFP measure to do policy analysis on productivity, we suggest looking for robust empirical IC results using several TFP measures. Ackerberg et al (2006) said;

> “*Finding that production function parameters are consistent across multiple techniques with different assumptions is surely more convincing than only using one*”.

For this purpose, we use eight productivity (TFP) measures that best fit with the characteristics of our data set: two levels of aggregation (restricted and unrestricted), with two parametric production functions (Cobb-Douglas and the Translog), with the Solow residuals for the two aggregation levels and applying Levinshon and Petrin (2003) (LP) and Ackerberg, Caves and Frazer (2006) (AC&F) structural estimators.

This paper aims to obtain robust IC partial effects on total factor productivity (TFP). TFP measures the effects of any other variable different from the inputs—labor (L), intermediate materials (M) and capital services (K)—, affecting the production (or sales) process. To be more specific, consider that the general production function is $Y_{it}=F(L_{it},M_{it},K_{it},TFP_{it};\alpha)$ and the productivity is indicated by TFP_{it}. The individual plants are indicated by the sub-index $i = 1, 2,..., N$, where N is the total number of plants in the sample and by the sub-index time $t = 1, 2, ..., T$, where T is the total number of years in the sample.[12]

The usual endogeneity of the inputs (L, M and K) is exemplified in the following simple Cob-Douglas production function model,

$$y_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + tfp_{it} \tag{1a}$$

$$tfp_{it} = \upsilon_{it} + e_{it} \tag{1b}$$

where $E(tfp_{it} / l_{it}, m_{it}, k_{it}, \alpha) \neq 0$, $E(e_{it} / l_{it}, m_{it}, k_{it}, \upsilon_{it}, \alpha) = 0$ and equation (2) is a regression model with unobservable explanatory variable; the unobserved productivity shocks (υ_{it})

$$y_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + \upsilon_{it} + e_{it} \tag{2}$$

and the input variables (L, M and K) are "exogenous" after conditioning on the unobservable productivity shocks.

Our TFP estimation procedure is justified on the following simplified *simultaneous equations model (SEM)*; (3a) production function (Cobb-Douglas in this case) and (3b) the IC determinants of the usually unobserved firm specific fixed-effects,

$$y_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + \upsilon_{it} + e_{it} \tag{3a}$$

[12] In the IC surveys, N is large and T is small. Lower letters indicate that the variable is in logarithms (logs).

$$\upsilon_{it} = \alpha'_{IC} IC_{Pi} + \alpha'_{C} C_{Pi} + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P + \zeta_{it} \tag{3b}$$

where IC_i and C_i of equation (3b) are *plant-level fixed effect vectors* of investment climate variables and other control variables, while D_j and D_t are vectors of industry (j) dummies and year (t) dummies, respectively. Notice that the time dummies (D_t) capture part of the momentum that productivity has and that is usually captured by considering a first order autocorrelation process (first order Markov condition in the structural model) with high persistence.[13]

The usual unobserved fixed effects, ($a_{Pi} = \alpha'_{IC} IC_{Pi} + \alpha'_{C} C_{Pi}$) included in the υ_{it} component of equation (3a) are proxy here by the set of observable fixed effects given by IC and C variables of (3b).

$$y_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + \alpha'_{IC} IC_{Pi} + \alpha'_{C} C_{Pi} + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P + \zeta_{it} + e_{it}. \tag{4}$$

Therefore, the extended production function (4) represents the *conditional expectation* plus a composite random error term that is the sum of the unpredictable productivity shocks (ζ_{it}) and the idiosincratic shocks (e_{it}),

$$\begin{aligned} E(y_{it} / l_{it}, m_{it}, k_{it}, IC_{P,i}, C_{P,i}, D_j, D_t, \theta) = \\ = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + \alpha'_{IC} IC_{Pi} + \alpha'_{C} C_{Pi} + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P \end{aligned}$$

the random error term $u_{it} = \zeta_{it} + e_{it}$, of the *extended production function* is assumed to be conditionally uncorrelated with the explanatory *L, M, K, IC, C* and dummy variables (D),

$$\begin{aligned} E\left[u_{it} / l_{it}, m_{it}, k_{it}, IC_{Pi}, C_{Pi}, D_j, D_t\right] = 0 \\ and \quad Var\left[u_{it} / l_{it}, m_{it}, k_{it}, IC_{Pi}, C_{Pi}, D_j, D_t\right] = \sigma^2_{u,it}. \end{aligned}$$

Notice that we condition on the observable fixed-effects, (IC_i) and (C_i), and on certain industry (D_j) and time dummies (D_t) to get the orthogonally condition of the inputs L, M

[13] We could do that but since most ICs of developing countries are very unbalanced, we prefer not to lose many observations (firms) when allowing for the AR(1) version of TFP. However, we plan to do that in the near future when having access to balanced panel based on investment climate surveys.

and K with the error term (u_{it}).[14] Without conditioning in IC and C variables there is correlation between the regression error and the inputs (L, M and K) coming from the *common causes* generated by the observable fixed IC and C effects[15].

The previous argument applies to other functional forms. For example consider the TRANSLOG extended production function,

$$\begin{aligned} y_{it} &= \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + \alpha_{LL}(l_{it})^2 + \alpha_{MM}(m_{it})^2 + \alpha_{KK}(k_{it})^2 + \\ &+ \alpha_{LM} l_{it} m_{it} + \alpha_{LK} l_{it} k_{it} + \alpha_{MK} m_{it} k_{it} + \alpha'_{IC} IC_i + \alpha'_C C_i + \alpha'_{Ds} D_j + \alpha'_{Dt} D_t + \alpha_P + u_{it} \end{aligned} \tag{3}$$

that is only a "local approximation" to the unknown function and therefore might not give very reliable globally parameter estimates.

The popular two-step approach to estimate IC partial effects could be used if the corresponding TFP measure is obtained from the nonparametric Solow´s residuals, Solow (1957), using accounting techniques based on the cost-shares of Hall (1990). In the first step, (4a), we get the estimation of TFP (Solow residual), where $\bar{s}_j$ is the average of the corresponding input cost-shares from the last two years.[16] In the second step, (4b), we estimate the partial effects of the IC variables on TFP,

$$\hat{tfp}_{it} = y_{it} - \bar{s}_L l_{it} - \bar{s}_M m_{it} - \bar{s}_K k_{it} \tag{4a}$$

$$\hat{tfp}_{it} = \alpha'_{IC} IC_i + \alpha'_C C_i + \alpha'_{Ds} D_j + \alpha'_{Dt} D_t + \alpha_P + w_{it} \tag{4b}$$

The advantage of using Solow´s residual, is that it does not require the inputs (L, M, K) to be exogenous nor the input-output elasticities to be constant. The drawback is that it requires having constant returns to scale (CRS) and at least competitive input markets.

---

[14] In all the regressions with different TFP measures, we always include 7 dummy variables (D_r, r = 1, 2, ..., 7) and a constant term (intercept). That is, we control for 8 industry effects (food and beverages, textiles, apparels, wood & furniture, paper&edition, chemicals, rubber&plastics, non-metallic products and machinery&equipment-Metallic products) and two years dummies for the three years of data as indicated by Table 1 of the appendix.

[15] In the empirical application we also consider standard errors of the parameters that are robust to clusters (region-industry) and/or to heteroskedasticity and autocorrelation (HAC).

[16] In the restricted case we have $\bar{s}_j = (1/2)(s_{jt} + s_{jt-1})$ for j = L, M and K, where s_j is the corresponding cost-share and the average runs across the whole sample. In the unrestricted case, the averages are computed industry by industry.

Other two-step procedures are applied when, in the first step, we estimate productivity using the L&P and AC&F algorithms.[17] In the case of the AC&F specification we use the one-step procedure proposed by Wooldridge (2009) due to the efficiency gains obtained relative to the standard AC&F procedure. Thus, in the empirical section we end up with eight different TFP estimation procedures of the IC effects on TFP.

2.1 Comparison of our Extended Production Function Approach with the Structural Production Function Approaches.

The structural methods to estimate production functions have gained popularity in the last years, beginning with Olley and Pakes (1996) (O&P for short), with key contributions of Levinshon and Petrin (2003) (L&P) and Ackerberg, Caves, and Frazer (2006) (AC&F) and Wooldridge (2009). In all the cases, estimation is based on using lagged input decisions as instruments, and share some similarities and divergences with the model proposed here. Since AC&F encompasses the main characteristics of O&P and L&P, we concentrate on this method to compare it with the model proposed here.

The structural model in AC&F can be written in this case as,[18]

$$y_{it} = \alpha_L l_{it} + \alpha_K k_{it} + \upsilon_{it} + e_{it} \quad (5a)$$

$$\upsilon_{it} = \Psi\left[\upsilon_{it-1}\right] + \alpha_P + \xi_{it} \quad (5b)$$

$$\upsilon_{it} = f_t^{-1}(m_{it}, k_{it}, l_{it}) \quad . \quad (5c)$$

The unobserved productivity (υ_{it}) corresponding to both production functions, (5b) and (1c), is "proxy" by two different approaches based on two different information sets. While AC&F requires having certain types of dynamic panel structure,[19] our approach

---

[17] In the case of the AC&F estimation we use the one-step modification proposed by Wooldridge (2009), which has been proved to be more efficient.

[18] See Appendix I for more details on the OP, LP and ACF models.

[19] The corresponding equation (10b), in the traditional *dynamic panel literature* of Chamberlain (1982), Anderson and Hsiao (1982), Arellano and Bond (1991) and Blundell and Bond (1998, 2000), is the following; $\upsilon_{it} = a_i + \alpha'_{Ds} D_j + \omega_{it} + \alpha_P + \xi_{it}$ where ω_{it} follows an AR(1) process, $\omega_{it} = \rho\omega_{it-1} + \varsigma_{it}$ where the AR(1) coefficient (ρ) is high and close to one (persistent productivity shocks) . In our approach this high AR(1) is proxy by a flexible deterministic trend with changing coefficients given by $\alpha'_{DT} D_t$.

can be used in simple cross sections or in dynamic panels with trending data but with uncorrelated errors. The corresponding extended production functions is;

$$y_{it} = \alpha_L l_{it} + \alpha_K k_{it} + \Psi\left[f_{t-1}^{-1}(m_{it-1}, k_{it-1}, l_{it-1})\right] + \alpha_P + \xi_{it} + e_{it} \tag{6}$$

However, following AC&F, equation (6) requires a two-step approach; first and estimate of $f^{-1}(.)$ and Ψ, using equations (5b) and (5c), and second a final estimate of the α_L and α_K from the sample analogue of the following two orthogonality conditions $E\left[\xi_{it}(\alpha_L, \alpha_K) / \begin{bmatrix} k_{it} \\ l_{it-1} \end{bmatrix}\right] = 0$. Notice that none of the structural TFP estimation procedures, O&P, L&P or AC&F, allow us to consider fixed effects in equation (6).

From now on, our *baseline model* to carry out the robustness tests will be the extended production function (2).[20] For this purpose, equations (2) and (6) can be nested into model (7).

$$y_{it} = \alpha_L l_{it} + \alpha_K k_{it} + \Psi\left[f_{t-1}^{-1}(m_{it-1}, k_{it-1}, l_{it-1})\right] + \alpha'_{IC} IC_{Pi} + \alpha'_C C_{Pi} + \alpha'_{Ds} D_j + \alpha'_{DT} D_t + \alpha_P + e_{it} \tag{7}$$

This nested model allow us to test whether our IC elasticity estimates from (2) differ by adding the cubic polynomial approximations suggested L&P procedure. The results will be discussed later on in the empirical section.

2.2 Strategies for IC Variable Selection

The econometric methodology applied for the selection of the IC and C variables goes from the general to the specific. Once we have a parsimonious model with only

---

We hope that The World Bank will create soon balanced dynamic panel of ICS so that we could evaluate the robustness of our empirical results.

[20] Heterogeneous and time varying input-output elasticities ($\alpha_{j,it}$) could be estimated by nonparametric procedures, index number techniques (Solow 1957, Diewert and Nakamura 2002) or estimated by regression techniques assuming that the input-output elasticity parameters are constant. In this paper, we will consider two options: a) the *unrestricted case* where constant input-output elasticities are considered to vary at the *industry level*, and b) the *restricted case* where the elasticity parameters are considered to be constant at the *aggregate level.*

significant variables, we test for omitted variables to make sure that we did not deleted relevant IC variables due to the strong multicolinearity of the initial general model.

The *omitted variables* problem that we encounter, starting from a too simple model generates biased and inconsistent parameter estimates. On the contrary, adding *irrelevant variables* (meaning starting from a very general model with some variables that are irrelevant) might suffer multicolinearity among IC variables providing unbiased and consistent but inefficient estimates. Therefore, we start from a general model, such as equations (4) with most for the 120 IC variables included at once, and we reduce this general model to a simpler one with only relevant (significant) variables[21]. We also start adding IC variables to our selected model to check if we had omitted a relevant IC variable in the process (specification test). Notice that the final estimated model is efficiently estimated once we have deleted insignificant or irrelevant variables.

Going from general-to-specific is usually recommended to avoid having omitted variable biases and spurious correlations, see Hendry and Nielsen (2007). Consider a regression with $\boldsymbol{n}$ irrelevant variables. Then the average number of variables found significant by chance at the α significant level is $\boldsymbol{n\alpha}$. Say α=0.05 and n=40 then $n\alpha$=2. That is, on average, 2 irrelevant variables are included and 38 variables are correctly excluded if *repeated t-test* are used. If α is reduced to α=0.01, as it is sometimes suggested when doing repeated t-testing, and n=40 then average number of variables found significant is reduced to $n\alpha$=0.4**.** However, the main problem of reducing the significance level $\boldsymbol{\alpha}$, is that we are also reducing the power of the t-test, making the detection of relevant variables difficult (which is a misspecification with crucial implications in terms of spurious correlations). Monte Carlo evidence shows from Hoover and Perez (1999) and Hendry and Krolzig (2001, 2005) that general-to-specific modeling has a small search cost; that is a small additional cost in terms of size and power that arise by doing *repeated testing with multiple path selection algorithms* starting in a general unrestricted model (GUM) that is not the true local DGP.

---

[21] Sometimes, in the final regression model, we leave IC variables that are not individually significant but are relevant for the model, either because they have a jointly significant affect with other variables or are significant in other TFP measures. When this happens it could be due to the presence of multicolinearity among some of the explanatory variables of the production function (Translog case specially) or among other IC variables.

In the reduction process we do not delete all insignificant variables at once, since due to multicolinearity, if we drop one variable that is highly correlated with others, some of the insignificant variables might become significant. An informative statistic for this purpose is the variation of the R^2 of the regression (or the variation in the standard error of the regression). The R^2 of the final simplified models, with only significant or relevant variables, are included in Table 8 for Costa Rica. Those R^2 of the reduced models are smaller but very close to the R^2 of the most general regression model we started with. We applied this iterative procedure, eliminating the least significant variables leaving, for interpretive purposes, at least one IC variable from each broad IC category (infrastructure, bureaucracy/corruption, crime, technology and quality, human capital, corporate governance, etc.). Once we have a reasonable parsimonious model we start testing for omitted IC variables block by block, to see if due to the multicolinearity among IC variables we deleted a relevant IC variable in the reduction process. Notice, that most automatic modeling process does not consider this steps going from the specific to the general step. The reason is that they are based on orthogonal regressors and therefore multicollinearity problems are not an issue.

2.3 Omitted IC Variables and Exogenous IC Variables

Taking region-industry averages in investment climate surveys is useful to avoid dropping a large number of establishments with missing relevant IC information that would create an important omitted variables problem (bias and inconsistent parameter estimates). In fact, a key aspect of our econometric methodology is that unobservable fixed effects are proxy by the 120 IC plant-level variables. However, without taking region-industry averages we would have ended up with less than 20% of the IC variables of the sample. The particular list of IC variables to be transformed as industry-region averages are based on two considerations: first, having a large number of missing values in those IC variables and, second, the possibility of being an endogenous variable. In our case, out of the 120 IC variables, 67 were used in industry-region form, while the remaining 53 IC variables were kept as plant-level variables.

There are, obviously, drawbacks associated with the use of the industry-region averages. Firstly, since we are using industry region averages, the interpretation of the partial effects should take into account this fact. The IC variables in average form should be interpreted as the overall investment climate conditions in the region and industry in which the firms operate. Secondly, the OLS estimators using the industry region averages will be less efficient than those using the plant level variables, provided the averages present lower variability.[22] Thirdly, one should be aware of multicolinearity problems, since the IC averages are likely to be highly correlated among them.[23] In order to avoid having a high degree of multicolinearity, and because we want to preserve the observable fixed effects interpretation of some of them, we never use all the IC variables in average region-industry form. For example, out of the 26 IC and C significant variables that we used in the final TFP model, 16 are used in average form, while keeping the plant level information for the remaining 10 IC and C variables.

An important econometric problem that we always have to face when estimating equations like (2) is that some of the IC and C explanatory variables might depend on firm´s decisions and therefore they could be endogenous. Examples of IC variables those state variables for the firm are for example: “sales declared to taxes”, “payments to obtain a contract with the government”, “firms belonging to a trade association” or “firms having access to a credit line”. The values of those IC variables can be modified by firm´s decisions. Therefore, some degree of correlation with the error term e_{it} of equation (2) could appear. In order o control for the endogeneity of these IC variables, we use two alternatives approaches: first, using region-industry IC variables for some of them and, second testing for the endogeneity (Hausman test) of IC variables.

When using *region-industry averages* ($\overline{IC}$) instead of the original variables, we have $\overline{IC}_{s,i} = \hat{\delta}_0 + \hat{\delta}_1' D_j + \hat{\delta}_2' D_R$, where D_j and D_R are vectors of industry and region dummy variables, and $(\hat{\delta}_0, \hat{\delta}_1', \hat{\delta}_2')$ are OLS parameter estimates. We assume that the industry region averages are exogenous in the short run period of three years, provided the location decisions of the firm are predetermined, i.e. decisions taken before the surveys were

[22] Nonetheless, even before this lost of efficiency, much of the IC partial effects remains significantly different from zero.
[23] In fact, we have 7 regions and 8 industries in total, so the maximum number of IC variables in average form that could be included is 56 to avoid perfect multicolinearity.

conducted, and now the "exogeneity" condition for consistency involves simply that two covariances are zero; $C(D_j, e_{it})= C(D_R, e_{it})=0$.[24]

However, since we are also using 10 plant level IC variables, one could argue that the endogeneity problem has not been fully addressed. In order to make sure that we do not have a serious endogeneity problem when using those IC variables at the plant level we use the Hausman test, considering the corresponding industry region average as excluded instruments of the plant level variables.[25] The results are satisfactory, as the Hausman test does not reveal any remaining endogeneity problem in the model; we do not reject the null of consistent OLS estimators with a p-value equal to 0.488.

Important causality questions related the degree of endogeneity of the production function inputs and IC variables, mainly due to simultaneity and reverse causality, will be addressed in the next section based on simulation exercises.

3. Monte Carlo Simulations: Omitted IC variables and Endogenous IC Variables

To evaluate this aspect of endogeneity of the inputs and IC variables in TFP regressions we did the following Monte Carlo simulation exercises[26]. For each i^{th}-firm of the population, the production in this economy is generated by the following data generation process (DGP), which is a system of equations,

---

[24] When some instruments and or regressors are estimated in a first stage the asymptotic variance needs to be adjusted because of the generated instruments, see Pagan (1984), Newey (1984), Murphy and Topel (1985) and Newey and McFadden (1994). More precisely, when testing the null hypotheses $H_0 : \alpha_j = 0$, (see equation 7), the usual test statistic has a limiting standard normal distribution under H_0. However, when $\alpha_j \neq 0$ standard t statistics will not be asymptotically valid and an adjustment is needed for the asymptotic variances of all estimators of generated regressors. A standard solution for this problem is to compute the bootstrap estimate of the standard errors of the estimated coefficients of (7). We have used this method without significant changes and the results are available upon from request.

[25] Since we have more industries and regions than candidates to be endogenous variables, the model is identified by 2SLS.

[26] See Escribano and Pena (2014), for a more detailed analysis based on Monte Carlo Simulations and for the algebraic derivations of the asymptotic bias in each TFP regression equation under different sources of model misspecification like; omitted variables, irrelevant variables, endogeneity and reverse causality.

$$y_i = \alpha_L l_i + \alpha_K k_i + tfp_i \qquad (8a)$$
$$tfp_i = \alpha_0 + \alpha_1 ic_{1i} + \alpha_2 ic_{2i} + e_{\omega i} \qquad (8b)$$
$$l_i = \beta_0 + \beta_1 ic_{1i} + \beta_2 ic_{2i} + \beta_3 wage_i + e_{li} \qquad (8c)$$
$$k_i = \gamma_0 + \gamma_1 ic_{1i} + \gamma_2 ic_{2i} + \gamma_3 r_i + e_{ki} \qquad (8d)$$
$$ic_{1i} = \delta_0 + \delta_1 D_i + \delta_2 ic_{2i} + e_{1i} \qquad (8e)$$
$$ic_{2i} = \rho_0 + \rho_1 tfp_i + \rho_2 ic_{3i} + e_{2i}. \qquad (8f)$$

The first equation is a Cobb-Douglas production function with productivity shocks (ω_i); the second equation is a classical labor demand equation; the third one is the demand for capital equation. Notice that *tfp* and the two demand equations, (8b)-(8d), depend on two important environmental variables related to the investment climate (ic_1 and ic_2). Assume, for simplicity that wages, the user cost of capital (r) and variables D and ic3 are all *exogenous variables*.

In terms of the *extended production function*, substituting (8b) in (8a) and (8a) in (8f), the system of equations becomes,

$$y_i = \alpha_L l_i + \alpha_K k_i + \alpha_0 + \alpha_1 ic_{1i} + \alpha_2 ic_{2i} + e_{\omega i} \qquad (9a)$$
$$l_i = \beta_0 + \beta_1 ic_{1i} + \beta_2 ic_{2i} + \beta_3 wages_i + e_{li} \qquad (9b)$$
$$k_i = \gamma_0 + \gamma_1 ic_{1i} + \gamma_2 ic_{2i} + \gamma_3 r_i + e_{ki} \qquad (9c)$$
$$ic_{1i} = \delta_0 + \delta_1 D_i + \delta_2 ic_{2i} + e_{1i} \qquad (9d)$$
$$ic_{2i} = \rho_0 + \rho_1 (y_i - \alpha_L l_i - \alpha_K k_i) + \rho_2 ic_{3i} + e_{2i}. \qquad (9e)$$

Notice that in equation (9e) there are two nonlinear parameter restrictions in the coefficients of l and k.

All the error terms (e_i) of the system of five equations, (9a)-(9e), are assumed to be independent, identically distributed and Gaussian with cero mean and variance-covariance the identity matrix, *i.i.d.*$N(0_5,I_5)$.

Inconsistent OLS parameter estimates of the input output elasticities of labor and capital of equation (9a) are obtained due to either/both: a) omitted variables, like ic_1 and ic_2 or b) simultaneous equation problems ($\rho_1 \neq 0$), generating in both cases correlation between TFP *(tfp)* and the inputs labor and capital (L and K); *endogenous* explanatory variables in both cases. However, in the extended production functions, equations (9a) and (11a), the error term e_ω is a *structural error term* and uncorrelated with the explanatory variables and we can infer *causal relations* as long as $\rho_1=0$ (no simultaneity) in (9e) or (11c).

This system of equations (9) is simplified by written it in terms of the *total factor productivity (tfp),* either by assuming that tfp is observable or that it is obtained from growth accounting techniques or the Solow´s residual. In this 2-step procedure, in the first-step the total factor productivity (tfp) is obtained as the Solow´s residual, based on standard growth accounting techniques, and in the second-step equation (11a) is estimated by OLS.

In this case, the five equation system, (9a) to (9e), is simplified to a three equations system *with i.i.d.N(0_3,I_3) error terms* and without the nonlinear parameter restrictions,

$$tfp_i = \alpha_0 + \alpha_1 ic_{1i} + \alpha_2 ic_{2i} + e_{\omega i} \qquad (10a)$$
$$ic_{1i} = \delta_0 + \delta_1 D_i + \delta_2 ic_{2i} + e_{1i} \qquad (10b)$$
$$ic_{2i} = \rho_0 + \rho_1 tfp_i + \rho_2 ic_{3i} + e_{2i}. \qquad (10c)$$

The purpose of the Monte Carlo simulations of this section is to evaluate the sources of inconsistencies and biases generated by OLS estimation of equations (8a), (8b) and (9a), under different assumptions and modeling misspecifications conditions. In particular we want to evaluate the OLS bias generated by two of the main endogeneity sources[27] we face in empirical applications; *i) simultaneity and ii) reverse causality.*

For simplicity of the algebraic derivations, and without los of generality, in the following *Monte Carlo simulations* we will assume that the labor is i.i.d. and concentrate on the inconsistencies on the coefficients of capital and IC variables. We concentrate on the evaluation of the OLS bias of the input-output elasticity of capital (α_k) and of the two investment climate effects (ic_1 and ic_2) on productivity (α_1 and α_2) under different modeling misspecifications.

The *linear projection* of the error term $e_{i\omega}$ onto the regressors of equation (9a) is given by $e_{i\omega} = b_0 + b_L l_i + b_K k_i + b_1 ic_{i1} + b_2 ic_{i2} + r_{ie\omega}$. Then equation (9a) could be writen in terms of orthogonal errors ($r_{ie\omega}$) as,

$$y_i = (\alpha_0 + b_0) + \alpha_L l_i + (\alpha_K + b_K)k_i + (\alpha_1 + b_1)ic_{i1} + (\alpha_2 + b_2)ic_{i2} + r_{ie\omega}. \qquad (11)$$

These b´s coefficients are useful to evaluate the desired asympotic bias of the coefficients of k, ic_1 and ic_2 and given by,

[27] Other usual misspecification cases like; omitted IC variables, irrelevant IC variables and proxy IC variables are studied in Escribano and Pena (2014).

$$p\lim(\hat{\alpha}_k-\alpha_k)=b_k=\frac{C(e_\omega,k)}{\sigma_k^2}-b_1\frac{C(ic_1,k)}{\sigma_k^2}-b_2\frac{C(ic_2,k)}{\sigma_k^2} \qquad (12a)$$

$$p\lim(\hat{\alpha}_1-\alpha_1)=b_1=\frac{C(e_\omega,ic_1)}{\sigma_{ic_1}^2}-b_k\frac{C(k,ic_1)}{\sigma_{ic_1}^2}-b_2\frac{C(ic_2,ic_1)}{\sigma_{ic_1}^2} \qquad (12b)$$

$$p\lim(\hat{\alpha}_2-\alpha_2)=b_2=\frac{C(e_\omega,ic_2)}{\sigma_{ic_2}^2}-b_k\frac{C(k,ic_2)}{\sigma_{ic_2}^2}-b_1\frac{C(ic_1,ic_2)}{\sigma_{ic_2}^2}. \qquad (12c)$$

Solving the system (12a)-(12c) we get,

$$\begin{pmatrix} p\lim(\hat{\alpha}_k-\alpha_k) \\ p\lim(\hat{\alpha}_1-\alpha_1) \\ p\lim(\hat{\alpha}_2-\alpha_2) \end{pmatrix}=\begin{pmatrix} b_k \\ b_1 \\ b_2 \end{pmatrix}=\begin{pmatrix} 1 & \frac{C(ic_1,k)}{\sigma_k^2} & \frac{C(ic_2,k)}{\sigma_k^2} \\ \frac{C(k,ic_1)}{\sigma_{ic_1}^2} & 1 & \frac{C(ic_2,ic_1)}{\sigma_{ic_1}^2} \\ \frac{C(k,ic_2)}{\sigma_{ic_2}^2} & \frac{C(ic_1,ic_2)}{\sigma_{ic_2}^2} & 1 \end{pmatrix}^{-1}\begin{pmatrix} \frac{C(e_\omega,k)}{\sigma_k^2} \\ \frac{C(e_\omega,ic_1)}{\sigma_{ic_1}^2} \\ \frac{C(e_\omega,ic_2)}{\sigma_{ic_2}^2} \end{pmatrix} \qquad (13)$$

Theorem 1 (Simultaneous Equations): Under the assumptions that the errors of the system of equations (9a)-(9e) are i.i.d.N(0_5, I_5), the exogenous variables (L, r, D and ic_3) are also i.i.d (for simplicity) then the *simultaneous equation bias generated by having $\rho_1 \neq 0$,* in equation (9e) affects all relevant piecewise covariances of the three main variables (k, ic_1 and ic_2) and therefore the OLS estimator of equation (9a) is inconsistent, asymptotically biased and *the biased depends on the noise to signal ratio of each explanatory variable*;

$$p\lim(\hat{\alpha}_k-\alpha_k)\neq 0, \qquad p\lim(\hat{\alpha}_1-\alpha_1)\neq 0, \quad \textit{and} \quad p\lim(\hat{\alpha}_2-\alpha_2)\neq 0.$$

$$\frac{C(e_\omega,k)}{\sigma_k^2}=\left(\frac{\rho_1\left(\gamma_1\delta_2+\gamma_2(1-\delta_2\alpha_1)\right)+\rho^2{}_1\left(-\gamma_1\delta_2\alpha_2+\gamma_2(\alpha_1\delta_2-\alpha_2)\right)}{1-\rho_1(\alpha_2+\delta_2\alpha_1)}\right)\frac{\sigma_{e\omega}^2}{\sigma_k^2} \qquad (14a)$$

$$\frac{C(e_\omega,ic_1)}{\sigma_{ic_1}^2}=\left(\frac{\rho_1\delta_2\left(1-\rho_1\alpha_2\right)}{1-\rho_1(\alpha_2+\delta_2\alpha_1)}\right)\frac{\sigma_{e\omega}^2}{\sigma_{ic_1}^2} \qquad (14b)$$

$$\frac{C(e_\omega,ic_2)}{\sigma_{ic_2}^2}=\rho_1\left(\frac{\rho_1\delta_2\alpha_1}{1-\rho_1(\alpha_2+\delta_2\alpha_1)}+1\right)\frac{\sigma_{e\omega}^2}{\sigma_{ic_2}^2} \qquad (14c)$$

Proof: The covariances are obtained from the properties and the algebra of linear projections. The inconsistency of the OLS estimator follows substituting the covariances (14a)-(14c) in (13).

Corollary 1 (Reverse Causality): Under the assumptions of Theorem 1, the simultaneous equation bias from having $\rho_1 \neq 0$ still affects all relevant piecewise covariances of the three main explanatory variables (k, ic_1 and ic_2) *even if ic_2 is an irrelevant* in equation (9a) with coefficient $\alpha_2=0$, the OLS estimator in (9a) is inconsistent and asymptotically biased;

$$p\lim(\hat{\alpha}_k - \alpha_k) \neq 0, \qquad p\lim(\hat{\alpha}_1 - \alpha_1) \neq 0, \quad and \quad p\lim(\hat{\alpha}_2) \neq 0.$$

$$\frac{C(e_\omega, k)}{\sigma_k^2} = \left(\frac{\rho_1\left(\gamma_1\delta_2 + \gamma_2(1-\delta_2\alpha_1)\right) + \rho^2{}_1\gamma_2\alpha_1\delta_2}{1-\rho_1\delta_2\alpha_1} \right) \frac{\sigma_{e\omega}^2}{\sigma_k^2} \tag{15a}$$

$$\frac{C(e_\omega, ic_1)}{\sigma_{ic_1}^2} = \left(\frac{\rho_1\delta_2}{1-\rho_1\delta_2\alpha_1} \right) \frac{\sigma_{e\omega}^2}{\sigma_{ic_1}^2} \tag{15b}$$

$$\frac{C(e_\omega, ic_2)}{\sigma_{ic_2}^2} = \rho_1 \left(\frac{\rho_1\delta_2\alpha_1}{1-\rho_1\delta_2\alpha_1} + 1 \right) \frac{\sigma_{e\omega}^2}{\sigma_{ic_2}^2} \tag{15c}$$

Proof: Immediate by making $\alpha_2=0$ in the covariances of Theorem 1.

Monte Carlo Simulations:

The two data generation process (DGP) considered are based on the following simultaneous equations system:

$$y_i = 0.6l_i + 0.4k_i + tfp_i \tag{16a}$$
$$tfp_i = 0.5 + 0.3ic_{1i} + \alpha_2 ic_{2i} + e_{\omega i} \tag{16b}$$
$$l_i = 5 + e_{li} \tag{16c}$$
$$k_i = 4 - 0.3ic_{1i} - 0.5r_i + e_{ki} \tag{16d}$$
$$ic_{1i} = 0.5 + 0.5D_i + 0.6ic_{2i} + e_{1i} \tag{16e}$$
$$ic_{2i} = 0.5 + \rho_1 tfp_i + 0.3ic_{3i} + e_{2i}. \tag{16f}$$

DGP(1): *Simultaneity without reverse causality*; equations (16a)-(16f) where $\alpha_2=0.3$ and $\rho_1=0.4$, (ic_2 cause tfp).

DGP(2): *Simultaneity with reverse causality*; equations (16a)-(16f) where $\alpha_2=0$ and $\rho_1=0.4$, (ic_2 do not cause tfp).

Under model uncertainty about the true DGP, nine regression models will be the estimated in the Monte Carlo simulations using Model(1) to Model(9). Under endogeneity in the ic_2 variable, generated in this case by having a simultaneous equation

($\rho_1 \neq 0$), we want to evaluate the corresponding bias in the coefficients of k, ic_1 and ic2 and in the coefficient of ic_3, a proxy variable for ic_2 variable. We will simulate the two most important causality issues (causality and reverse causality) we face in practice when estimating productivity impacts of economic variables, here investment climate variables, ic_{ji} variables from j=1,2 and 3.

Alternative empirical models estimated by OLS:

Production Function (PF) (simple PF Figures 1 and 2)

$$y_i = \alpha_0 + \alpha_L l_i + \alpha_K k_i + t\tilde{f}p_i \qquad Model(1)$$

Fully Extended Production Function **(**Fully-ext PF in Figures 1 and 2)

$$y_i = \alpha_0 + \alpha_L l_i + \alpha_K k_i + \alpha_1 ic_{1i} + \alpha_2 ic_{2i} + e_{\omega i} \qquad Model(2)$$

Semi-Extended Production Function (Semi-Ext PF in Figures 1 and 2)

$$y_i = \alpha_0 + \alpha_L l_i + \alpha_K k_i + \alpha_1 ic_{1i} + u_i \qquad Model(3)$$

Approximate Fully Extended PF (Approx Fully-ext PF in Figures 1 and 2)

$$y_i = \alpha_0 + \alpha_L l_i + \alpha_K k_i + \alpha_1 ic_{1i} + \alpha_3 ic_{3i} + \eta_i \qquad Model(4)$$

Fully Extended PF with Irrelevant variables (Irrelevant vars PF in Figure 1 and 2)

$$y_i = \alpha_0 + \alpha_L l_i + \alpha_K k_i + \alpha_1 ic_{1i} + \alpha_2 ic_{2i} + \alpha_3 ic_{3i} + v_i \qquad Model(5)$$

TFP equation corresponding to PF model (2)

$$t\tilde{f}p_i = \alpha_0 + \alpha_1 ic_{1i} + \alpha_2 ic_{2i} + e_{\omega i} \qquad Model(6)$$

TFP equation corresponding to PF model (3)

$$t\tilde{f}p_i = \alpha_0 + \alpha_1 ic_{1i} + u_i \qquad Model(7)$$

TFP equation corresponding to PF model (4)

$$t\tilde{f}p_i = \alpha_0 + \alpha_1 ic_{1i} + \alpha_3 ic_{3i} + \eta_{ii} \qquad Model(8)$$

TFP equation corresponding to PF model (5)

$$t\tilde{f}p_i = \alpha_0 + \alpha_1 ic_{1i} + \alpha_2 ic_{2i} + \alpha_3 ic_{3i} + \eta_{ii} \qquad Model(9)$$

4. Empirical Results on the impact of the IC variables on firms' Productivity

As we have pointed out in the previous section the robustness of these empirical results across TFP measures allows us to obtain *robust economic evaluations* of the IC effects of productivity which was the main purpose of this paper. So the natural question arising at this point is, what can we learn from this analysis about the investment climate conditions faced by Costa Rican firms?

4.1 Marginal Effects: Impacts on TFP of changes in IC variables

First, from the IC elasticities and semi-elasticities reported in Table 1,[28] we have 26 significant IC variables, 4 in the block of Infrastructure, 5 within Red tape, corruption and crime, 5 in Finance, 7 in the block of Quality, innovation and labor skills, and 5 in the group of Other control variables. The interpretation of the effects is done, as usual, in *ceteris paribus* terms, so the statement "*for firms facing the same investment climate conditions, the same input levels and operating in the same industry and year*" is what matters. Also, it is important to keep in mind that the interpretation of the partial effects when the IC variables are measured as industry –region averages is slightly different than in the usual case when we use the plant-level variable. In fact, a change in the industry-region averages can be thought of as an improvement in the overall investment climate conditions.

Thus, for firms operating in the same investment climate conditions and keeping everything else constant, decreasing the average (by region and industry) time that firms waste to clear customs by 1% could increase on average firm level productivity by

[28] The economic interpretation of each investment climate coefficient is contingent on the units of measurement of each IC variable and on the transformations performed on them (logs, fractions, percentages, qualitative constructions, etc.). Since all the productivity (TFP) measures considered here are always in logs, when the IC variable is also expressed in logs the estimated coefficient measures the constant *IC elasticity on TFP*. When the IC variable is not expressed in logs and is not a binary variable, the estimated coefficient is usually described as the *IC semi-elasticity on TFP.* While it is sometimes natural to express an IC variable in logs, for some type of IC variables it is more appropriate not to do so. For example, when an IC variables is a fraction or a percentage number with some data equal to 0 or close to 0. Notice however that expressing IC variables as fractions allow us to interpret also their coefficients as constant elasticities and not as semi-elasticities.

0.076%. Likewise, still in the Infrastructure block, if we reduce the average duration of the power outages, the average number of water outages, and the average time waiting to obtain and electricity supply, productivity could increase by 0.029, 0.217 and 0.128% respectively. In summary, successive improvements of the quality of the infrastructure used by the firms to generate output could lead to substantial improvements of the productivity levels.

Regarding the second block, Red tape, corruption and crime, the average percentage of sales declared to IRS for tax purposes, a measure of the degree of informality of the firms, has a positive effect on productivity, increasing the sales reported by 1% increases productivity by 1% too, *ceteris paribus*. The average number of inspections received has a harmful effect on productivity, in fact, decreasing it by 1% could increase productivity by 0.32%. The next variable payments to get a contract with the government is, in origin, a binary variable measuring the degree of corruption, in its average form it measures the proportion of firms (by industry and region) offering this payments. The effect is positive, meaning that increasing that proportion firm level productivity could increase. Finally, within this group, the average percentage of firms' sales never repaid and the number of days of work lost due to absenteeism have both a negative effect on productivity. In summary, the relevant IC variables of this group have to do with informalities, red tape and corruption.

In what refers to the financing sources of the firms, it is clear that having access to proper financing boosts firm level productivity, as the results show. Thus, those firms with access to a credit line are a 27.6% more productive than other firms. Similarly, the proportion of firms belonging to a trade association has a positive causal effect on TFP. Also, the profits as a percentage of total sales seem to spurs productivity.

Quality issues are statistically associated with productivity as the results show, pointing to a gap of 30.1% in terms of productivity between those firms with an ISO quality certification and those that do not. Innovation is also represented in this block of variables by the proportion of firms having a *New technological license*. More variables related with innovation, in this case embodied in the production process are the number of workers dealing with design and engineering issues and the percentage of the staff using computer controlled machinery and/or computers also have all a positive effects on productivity. Related to labor relations, the effect of the percentage of immigrant workers is negative, while the effect of training received by the staff is positive.

Although not considered a proper block of investment climate variables, the group of Other control variables, also reveals rather interesting relationships. For instance, we find that foreign direct investment, the number of competitors the percentage of capacity utilization and the imports have all positive effects on productivity.

4.2 Average Effects: IC Average Contributions to Average TFP

While the IC partial effects provide a useful view on the investment climate conditions of a given country, we can go one step further and evaluate the investment climate conditions of the average firm in the population. This exercise allows us to assess whether a given investment climate factor is important or not by considering the number of firms suffering that bottleneck or the magnitude of the problem firms have to deal with (e.g. number of power outages firs suffer on average). Thus, a given investment climate factor with a low partial effect can become important if most of the firms suffer that problem.

In particular, let the estimated extended Cobb-Douglas equation (2) be given by

$$y_{it} = \hat{\alpha}_1 l_{it} + \hat{\alpha}_2 m_{it} + \hat{\alpha}_3 k_{it} + \hat{\alpha}_1^{IC} IC_{1,i} + \ldots + \hat{\alpha}_r^{IC} IC_{r,i} + \hat{\alpha}'_C C_i + \hat{\alpha}'_{Ds} D_j + \hat{\alpha}'_{Dt} D_t + \hat{\alpha}_P + \hat{e}_{it} \quad (17)$$

and let the centered productivity (excluding the constant term) be given by $\hat{tfp}_{it}^d = y_{it} - \hat{\alpha}_L l_{it} - \hat{\alpha}_M m_{it} - \hat{\alpha}_K k_{it} - \hat{\alpha}_P$. Then the percentage mean relative IC contributions to the centered TFP is thus given by,

$$100 = \frac{\hat{\alpha}_1^{IC}\overline{IC}_1 + \ldots + \hat{\alpha}_r^{IC}\overline{IC}_r + \hat{\alpha}'_C\overline{C} + \hat{\alpha}'_{Ds}\overline{D} + \hat{\alpha}'_{Dt}\overline{D} + \bar{\hat{e}}}{\overline{\hat{tfp}^d}} 100 .$$

The problem with this measure is that some IC variables have positive effects and others IC variables might have negative effects. Therefore it might seems that the investment climate is nor relevant for TFP, because the positive values are compensated by the negative ones. Therefore, we prefer to consider the mean IC contributions in absolute value. That is,

$$100 = \frac{\left|\hat{\alpha}_1^{IC}\overline{IC}_1\right| + ... + \left|\hat{\alpha}_r^{IC}\overline{IC}_r\right| + \left|\hat{\alpha}'_C\overline{C}\right|}{\left|\overline{\hat{tfp}^d}\right|}100 + \frac{\left|\hat{\alpha}'_{Ds}\overline{D}\right| + \left|\hat{\alpha}'_{Dt}\overline{D}\right| + \left|\overline{\hat{e}}\right|}{\left|\overline{\hat{tfp}^d}\right|}100 \tag{18}$$

where $\left|\overline{\hat{tfp}^d}\right| = \left|\hat{\alpha}_1^{IC}\overline{IC}_1\right| + ... + \left|\hat{\alpha}_r^{IC}\overline{IC}_r\right| + \left|\hat{\alpha}'_C\overline{C}\right| + \left|\hat{\alpha}'_{Ds}\overline{D}\right| + \left|\hat{\alpha}'_{Dt}\overline{D}\right| + \left|\overline{\hat{e}}\right|$.

To compare the relative absolute contribution of each IC block over the total TFP associated only with IC variables, we compute the following IC percentage, relative to the demeaned TFP concept, in absolute value;

$$100 = \frac{\left|\hat{\alpha}'_I\overline{IC}_I\right| + \left|\hat{\alpha}'_R\overline{IC}_R\right| + \left|\hat{\alpha}'_F\overline{IC}_F\right| + \left|\hat{\alpha}'_Q\overline{IC}_Q\right| + \left|\hat{\alpha}'_C\overline{C}\right|}{\left|\overline{\hat{tfp}^D}\right|}100 \tag{19}$$

where $\hat{tfp}_{it}^D = \hat{tfp}_{it}^d - \hat{\alpha}'_{Ds}D_j - \hat{\alpha}'_{Dt}D_t - \hat{e}_{it}$, is the **demeaned TFP** concept only associated to the IC and C variables and $\left|\overline{\hat{tfp}^D}\right| = \left|\hat{\alpha}'_I\overline{IC}_I\right| + \left|\hat{\alpha}'_R\overline{IC}_R\right| + \left|\hat{\alpha}'_F\overline{IC}_F\right| + \left|\hat{\alpha}'_Q\overline{IC}_Q\right| + \left|\hat{\alpha}'_C\overline{C}\right|$. Notice that IC_I, IC_R, IC_F and IC_Q are the vectors containing the IC variables from the five blocks of investment climate variables described in Table B.2 of Appendix B; Infrastructure (IC_I), Red tape, corruption and crime (IC_R), Finance (IC_F) and Quality, innovation and labor skills (IC_Q) respectively.

The results of equations (18) and (19) are shown in Figures 2 and 3 of the appendix. From Figure 2 and following equation (19), the relative importance of the investment climate on explaining average TFP is of 27.7%, which in turn can be decomposed by blocks of investment climate variables (see right panel of Figure 2). Clearly the block of Red tape, corruption and crime variables dominates with a relative contribution of 34.4%, followed by Finance and Infrastructure.

Finally, Figure 3 shows the individual relative contributions of the IC variables. The most important one is the sales declared to IRS for tax purposes, a measure of formality. Also important are the number of inspections, time to get an electric connection, the profit as a percentage of sales, and the number of competitors in firms' main market.

The final TFP estimated coefficients of IC variables reported in Table 8 of the Appendix were obtained using this modeling selection strategy in all the countries. We included the set of IC variables that were *significant in at least one of the eight specifications.* The detailed empirical results are discussed in the next sections.

The estimations of the extended restricted production function, baseline model (3) are shown in Table 4 where we have ended up with 26 IC variables in the equation. The final estimated input-output elasticities are 0.312 for labor, 0.534 for materials and 0.121 for capital, all of them significant at 1% and the usual levels. Notice that, we cannot reject the null hypothesis of constant returns to scale (CRS) with a p-value equal to 0.345 (see last row of Table 4). In all the cases, inference is based on robust standard errors correcting for cluster correlation by industry and region.[29]

Remember that out of the 26 IC variables 16 are used in industry-region form. However, there are other 10 plant level IC variables for which we test if they are exogenous by carrying out Hausman tests. The null hypothesis of exogeneity of these 10 variables is not rejected with a p-value equal to 0.429. The instruments used for those plant level IC variables are the corresponding industry-region averages.

5. Testing the Robustness of the Productivity Analysis

The next question is whether this set of parameter estimates are robust across the eight specifications proposed is addressed in subsection 4.1. Later on in subsection 4.2 we compare the densities and cumulative distributions of the alternative TFP measures.

5.1 Testing the Robustness of the IC effects on TFP

In the second column of Table 1 we have the minimum and maximum parameter estimates across the eight specifications. In all the cases but one there are no change of signs and the numbers moves within a reasonable range of values. For instance, the elasticity of the number of *Days to clear customs to export* (IC variable #1, measured in logs) goes from -0.089, in the unrestricted Solow residual, to the -0.011 obtained by means of the LP procedure. The unique change in sign is in the variable *Payments to obtain a contract with the government* (var. #7, measured in percentage), in all the specifications the parameter associated to this variable have been estimated to be positive,

[29] We also applied the Newey-West estimator (HAC) to correct for heteroskedasticity and autocorrelation in the residuals. However the results are similar to those shown in Table 4. We present the cluster standard errors he advantage in Table 4 since they are more conservative than those obtained from Newey-West estimator.

in between 0.492 and 0.229, but in one case, the unrestricted Translog, however the effect is found to be statistically not different from zero in this estimation.

Notwithstanding the numerical similarity of the numerical coefficient estimates, found across the different specifications, is not a formal test of the robustness of our estimates. For a formal analysis we can test the null hypotheses that the parameters are all equal across specifications. That is, suppose we want to test whether the IC parameters are statistically the same in the *restricted Cobb-Douglas* and in the *restricted Solow* residual specifications and we have r IC variables in each equation, so they could be written respectively as

$$y_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + \alpha_1^{IC} IC_{1,i} + \ldots + \alpha_r^{IC} IC_{r,i} + \alpha_C C_i + \alpha_{Ds} D_j + \alpha_{Dt} D_t + \alpha_P + e_{it} \quad \text{(20a)}$$

$$\hat{tfp}_{it} = \gamma_1^{IC} IC_{1,i} + \ldots + \gamma_r^{IC} IC_{r,i} + \gamma_C C_i + \gamma_{Ds} D_j + \gamma_{Dt} D_t + \gamma_P + e_{it}^* \quad \text{(20b)}$$

for i=1, …,n. For simplicity of the notation C_i, D_j and D_t now represent scalars instead of vectors. A way to test the null $H_0 : \alpha_1^{IC} = \gamma_1^{IC}$ is to nest both models into a single one as in Fisher (1970). Let z_{it} be a new variable defined after stacking both y_{it} and $\hat{tfp}_{it}$ as

$$z_{it} = \begin{cases} y_{it} & \text{if } i = 1,\ldots.,n \\ \hat{tfp}_{it} & \text{if } i = n+1,\ldots.,2n \end{cases}$$

and define d_{it} as

$$d_{it} = \begin{cases} 0 & \text{if } i = 1,\ldots.,n \\ 1 & \text{if } i = n+1,\ldots.,2n \end{cases}$$

The new extended model (11), nest the two competing equations (10a) and (10b);

$$\begin{aligned} \mathrm{z}_{it} = {} & \alpha_L l_{it}^* + \alpha_M m_{it}^* + \alpha_K k_{it}^* + \alpha_1^{IC} IC_{1,i} + \ldots + \alpha_r^{IC} IC_{r,i} + \varphi_1^{IC}(d_{it} IC_{1,i}) + \ldots + \varphi_r^{IC} IC_{r,i} \\ & + \varphi_r^{IC}(d_{it} IC_{r,i}) + \alpha_C C_i + \varphi_C(d_{it} C_i) + \alpha_{Ds} D_j + \varphi_{Ds}(d_{it} D_j) + \alpha_{Dt} D_t + \varphi_{Dt}(d_{it} D_t) \quad (21) \\ & + \alpha_P + \varphi_P d_{it} + u_{it}^* \qquad i = 1,\ldots 2n \end{aligned}$$

where

$$l_{it}^{*}=\begin{cases} l_{it} & \text{if } i=1,....,n \\ 0 & \text{if } i=n+1,....,2n \end{cases},\quad m_{it}^{*}=\begin{cases} m_{it} & \text{if } i=1,....,n \\ 0 & \text{if } i=n+1,....,2n \end{cases},\quad k_{it}^{*}=\begin{cases} k_{it} & \text{if } i=1,....,n \\ 0 & \text{if } i=n+1,....,2n \end{cases}, \text{ and}$$

$$u_{it}^{*}=\begin{cases} u_{it} & \text{if } i=1,....,n \\ u'_{it} & \text{if } i=n+1,....,2n \end{cases}.$$

Now, testing $H_0:\varphi_1^{IC}=0$ is equivalent to test whether the difference between parameters is zero or $H_0:\alpha_1^{IC}-\gamma_1^{IC}=0$.[30] This procedure is repeated for each one of the eight specifications but always keeping as the model of reference the restricted Cobb-Douglas. That is, we are testing whether the IC parameters of the 8 alternative specifications are statistically equal to those of the restricted Cobb-Douglas.

Table 2 also shows the coefficient estimates of the 8 specifications, i.e. Cobb-Douglas, Translog, Solow Residual, L&P and AC&F, along with the p-value of the significance tests of $H_0:\varphi_r^{IC}=0$.[31] The results of these tests for the 26 IC variables included in the analysis are summarized in the third column of Table 1. For each variable, the number of statistically equal parameters across the eight specifications, along with the percentage in parentheses, is shown. The results are rather satisfactory; the effect of the *number of competitors* is statistically equal for all the specifications. *Days to clear customs to export* (#1) exhibits a statistically equal coefficient in 75% of the cases, the same occurs with other 7 variables. Five of the remaining variables show a statistically equal coefficient in 7 models. For another 7 variables the test reveals that the parameters are the same in 5 models.

More insights on the robustness of the parameter estimates are provided in Table 3. First, there is homogeneity in the directions (signs) of the IC effects across different specifications as compared to the restricted Solow residual, but in the case of one variable in the unrestricted Translog, but in that case the parameter is not statistically different

[30] Note that this test imposes the restriction that both e_{it} and e^*_{it} have the same variance. An equivalent alternative but without imposing such restriction is to define z_{it} as

$$z_{it}=(y_{it}-\hat{\alpha}_L l_{it}-\hat{\alpha}_M m_{it}-\hat{\alpha}_K k_{it})-\hat{tfp}_{it}$$

where $\hat{\alpha}_J$ are the OLS estimators of input-output elasticities in the extended Cobb-Douglas. Now the equivalent test can be carried out from the following regression

$$z_{it}=\rho_1^{IC}ic_{1,i}+...+\rho_r^{IC}ic_{r,i}+\rho_C C_i+\rho_{Ds}D_j+\rho_{Dt}D_t+\rho_P+e_{it}$$

where null hypothesis is now H_0: $\rho_1=0$. We have done these tests obtaining the similar results, so the restriction of equal variances of the residuals is not that influential. We thank Jean Marie Dofour's for suggesting this alternative procedure.

[31] The tests for the unrestricted estimations are available upon from request.

from zero. Second, out of the 26 IC variables 19 (73.1%) are statistically equal in the unrestricted Cobb-Douglas and restricted Translog with respect to the restricted Cobb-Douglas. Likewise, 18 variables have statistically equal parameters in the restricted Solow residual, 17 in the unrestricted Solow residual, 16 in the restricted Translog, 15 in the ACF specification and 14 in the LP case.

Tables 1, 2 and 3 demonstrate that the estimators of the different IC effects on firm level productivity are all reasonably robust across specifications. Based on these results we concentrate on only one set of parameters, say those coming from the restricted Cobb-Douglas, our baseline specification, to analyze and assess the importance of the investment climate on firms' productivity in Costa Rica, which is the aim of the next section. In summary, the key issue for getting this robustness is to control for the relevant IC firm-level information, avoiding omitting variables. Even if the analyst is only interested on the effects of say infrastructure on TFP, we have to control for the relevant IC information that we have on the other IC blocks (red tape, corruptions and crime, finance, quality, innovation and labor skills, etc.).[32]

5.2 Testing for Robustness Measuring TFP measures

Additional insight on the robustness of the results obtained is provided by Table 7 in which we show the main moments of the distributions of the eight productivity measures estimated, along with the Kolmogorov-Smirnov tests of equality of distributions. We consider the plain productivity measures obtained as $\hat{tfp}_{it} = y_{it} - \hat{\alpha}_L l_{it} - \hat{\alpha}_M m_{it} - \hat{\alpha}_K k_{it}$, and the productivity measure without the constant term, that is, $\hat{tfp}^d_{it} = \hat{tfp}_{it} - \hat{\alpha}_P = y_{it} - \hat{\alpha}_L l_{it} - \hat{\alpha}_M m_{it} - \hat{\alpha}_K k_{it} - \hat{\alpha}_P$. The reason of including this second productivity measure is to consider only the idiosyncratic firm-level productivity shock, isolating the constant technical efficiency, likely to ben he main source of inconsistencies among TFP measures.[33] Likewise, in Figure 1 we include the estimators of the productivity distributions and cumulative distributions.

[32] This was the reason why different units of the World Bank working with the same data set were obtaining different signs and magnitudes on common IC coefficients.

[33] We include the productivity measures without the constant term because the constant is likely to be estimated inconsistently, since the exogeneity conditions proposed only affects the parameters of the inputs, IC, C and D variables.

The question is whether we are able to obtain robust productivity measures. The answer is yes, especially when we exclude the constant term (or constant technical efficiency term). The Kolmogorov-Smirnov tests do not reject the null that the productivities are drawn from the same distribution in all the cases but when we consider heterogeneity with the unrestricted production functions. The graphical analysis also supports the robustness of the productivity measures obtained.

6. Conclusions

The investment climate is understood as the set of location specific factors shaping firms decisions of production. In this paper we have classified the investment climate in four blocks (Infrastructure, Red tape, Corruption and Crime, Finance and Quality, Innovation and Labor Skills) and we have proposed a robust general to specific methodology to estimate the partial effects of more than 100 of these investment climate variables on the productivity of Costa Rica´s firms. In particular we have proposed a baseline model consisting of an extended Cobb-Douglas production function using the information on the investment climate and other control variables to control for the usual transmission bias of OLS estimators. We refer to this model along the paper as *Restricted Cobb-Douglas* production function.

While developing this methodology was the first objective of the paper, the second one was to obtain robust partial effects of the investment climate variables on alternative measures of TFP. Our estimates produced robust impact evaluation across: 1) *different functional forms* of the production functions, 2) *different productivity measures* and 3) *different levels of aggregation* of the *input-output elasticities* (industry and country level).

We have shown that we are able to get all the expected signs on all IC coefficients and we have also proposed a test of equality of coefficients which allows us to conclude that the estimates show a reasonable robustness across specifications. We also showed that the TFP densities estimated are similar among them. Therefore, based on this robustness, based on a sensitivity analysis, in the further evaluation of the IC conditions in Costa Rica we only concentrate in the set of results obtained from our baseline model: the Restricted Cobb-Douglas production function.[34]

From the analysis we have found interesting relations between productivity and the investment climate, from which we can conclude that successive improvements of the investment climate conditions could lead to further productivity gains in Costa Rica. Besides the partial marginal effects of the IC we have proposed to evaluate the investment climate conditions on the average productivity of Costa Rica with important differences in terms of economic implications. The block of Red Tape, Corruption and Crime variables is the most relevant one for productivity, with main individual contributions from the degree of formality of the firms and the number of inspections received by the firms. Infrastructure is also important, for example decreasing the average time to get an electricity connection would spur productivity as the reduction of the prevalence of power outages or the time wasted in customs would do. Finally, granting firms proper financial resources and enhancing innovative activity would also produce further productivity gains.

[34] A different question is how to make cross-country comparisons based on TFP measures without comparing apples and oranges. To overcome this problem we have suggested the concept of demeaned TFP but deeper analysis of its properties is out of the scope of this paper.

References

Ackerberg, Daniel & Lanier Benkard, C. & Berry, Steven & Pakes, Ariel, 2007. "**Econometric Tools for Analyzing Market Outcomes**," Handbook of Econometrics, in: J.J. Heckman & E.E. Leamer (ed.), Handbook of Econometrics, edition 1, volume 6, chapter 63 Elsevier.

Ackerberg, D., K. Caves, and G. Frazer (2006). "Structural Identification of Production Functions". Mimeo.

Alexander, W. Robert J., Bell, John D., and Knowles, Stephen (2004). "Quantifying Compliance Costs of Small Businesses in New Zealand". Discussion paper, University of Otahgo. http://www.business.otago.ac.nz/econ/research/discussionpapers/DPO406.pdf.

Arellano M. and S. Bond (1991). "Some test of specification fro panel data: Monte Carlo evidence and application to employment equations". *The Review of Economics Studies*, 58, 277-297.

Chamberlain G. (1982). "Multivariate Regressions Models for Panel Data". *Journal of Econometrics* 18.

Barro R. J. and X. Sala-i-Martin (2004). Economic Growth. The MIT Press, second edition, Cambridge, Massachusetts.

Blundell R. and S. Bond (1998). "Initial conditions and moment restrictions in dynamic panel data models". *Journal of Econometrics*. 87, 115-143.

Blundell R. and S. Bond (2000). "GMM Estimation with Persistent Panel Data: An Application to Production Functions". *Econometric Reviews*, 19 (3), 321-340.

Bosworth, Barry and Susan Collins (2003). "The Empirics of Growth: An Update". The Brookings Institution. Washington, D.C. Processed.

Cole, H. L., L. E. Ohanian, A. Riascos and J. A. Schmitz Jr. (2004). "Latin America in the Rearview Mirror". National Bureau of Economic Research WP #11008, December.

Dethier J.J., M. Hirn and S. Straub (2008). "Explaining enterprise performance in developing countries with business climate survey data". *The World Bank Policy Research Working Paper* #4792.

Diewert W. Erwin and Alice O. Nakamura (2002) "The Measurement of Aggregate Total Factor Productivity Growth". J. J. Heckman and E. E. Leamer (eds.). Handbook of Econometrics, Vol. 6, forthcoming.

Djankov, Simeon, La Porta, Rafael, Lopez-de-Silanis, Florencio, and Shleifer, Andrei (2002). "The Regulation of Entry". *Quarterly Journal of Economics* 117, February 2002. 1-37.

Escribano Alvaro and J. Luis Guasch (2005). "Assessing the Impact of the Investment Climate on Productivity using Firm Level Data: Methodology and the Cases of Guatemala, Honduras and Nicaragua". Policy Research Working Paper # 3621, The World Bank, June.

Escribano Alvaro and J. Luis Guasch (2004). "Econometric Methodology for Investment Climate Assessments (ICA) on Productivity using Firm Level Data: The Case of Guatemala, Honduras and Nicaragua". Mimeo World Bank, June.

Foster L., J. Haltiwanger and C.J. Krizan (1998). "Aggregate Productivity Growth: Lessons from Microeconomic Evidence". *NBER Working Paper* W6803.

Hall, R. E. (1990). "Invariance Properties of Solow´s Productivity Residual". In Peter Diamond (ed.). Growth, Productivity, Employment. Cambridge: MIT Press, 1-53.

Hall, R. E. and C.I. Jones (1999). “Why Do Some Countries So Much More Output per Worker than Others?” *The Quarterly Journal of Economics*, 114, 1, 83-116.
Haltiwanger, John (2002). “Understanding Economic Growth”. The Need for Micro Evidence”. *New Zealand Economic Papers* 36 (1), 33-58.
He, Kathy S., Morck, Randall, and Yeung, Bernard (2003). “Corporate Stability and Economic Growth”. William Davidson Working Paper No. 553.
Hendry D. F. and H.M. Krolzig (2001). Automatic Econometric Model Selection. London: Timberlake Consultants Press.
Hendry D. F. and H.M. Krolzig (2005). “The properties of automatic Gets modeling”. *Economic Journal*, 115, C32-C61.
Hendry D. F. and B. Nielsen (2007). Econometric Modeling: A Likelihood Approach. Princeton University Press.
Hoover K.D. and S.J. Perez (1999). “Data mining reconsidered: Encompassing and the general-to-specific approach to specification search. *Econometrics Journal*, 2, 167-191.
Hulten Ch. R. (2001). “Total Factor Productivity: A Short Biography”. In Ch. R. Hulten, E.R. Dean and M. J. Harper (eds.) New Developments in Productivity Analysis, The University of Chicago Press, 1-47.
Jorgenson D.W. (2001). “Information Technology and the U.S. Economy,” *The American Economic Review*, Vol. 91, 1, 1-32.
Jorgenson D.W., F. Gollop and B. Fraumeni (1987). “Productivity and U.S. Economic Growth,” Cambridge: Harvard University Press.
Levinsohn J. and A. Petrin (2003). “Estimating Production Functions Using Inputs to Control for Unobservables”. *Review of Economic Studies*, 70, 317-341.
McMillan, John, (1998), “ Managing Economic Change: Lessons from New Zeland”, *The World Economy 21,* 827-43.
McMillan, John (2004), “A Flexible Economy? Entrepeneurship and Productivity in New Zeland”, Working Paper, Graduate Scholl of Business, Stanford University, Stanford, CA.
Olley G. S. and A. Pakes (1996). “The Dynamics of Productivity in the Telecommunications Equipment Industry”. *Econometrica*, Vol. 64, 6, 1263-1297.
Organization for Economic Cooperation and Development (2001). *Businesses' Views on Red Tape,* Paris, OECD. http://www1.oecd.org/publications/e-book/4201101E.PDF.
Prescot, Edward C. (1998) "Needed: A Theory of Total Factor Productivity". International Economic Review, 39, 525-552.
Rodrik, Dani, and Arvind Subramanian (2004). “From ‘Hindu Growth to Productivity Surge: The Mystery of the Indian Growth Transition”. Harvard University, Cambridge, Mass. Processed.
Solow R. M. (1957). “Technical Change and the Aggregate Production Function” *The Review of Economics and Statistics*, 39 (3), 312-320.
Wilkinson, Bryce (2001). “Constraining Government Regulation”. NZ Business Roundtable. http://www.nzbr.org.nz/documents/publications/publications-2001/constraining_govt.pdf.
Wooldridge J. M. (2002). Econometric Analysis of Cross Section and Panel Data. The MIT Press. Cambridge, Massachusetts.
Wooldridge J. M. (2009).”On estimating firm-level production functions using proxy variables to control for unobservables”. Economics Letters, Volume 104, Issue 3, September 2009, Pages 112–114.

World Bank 2003. "Doing Business in 2004" Understanding Regulation. Washington, D.C. World Bank.

__________ 2004a. "Doing Business in 2005: Removing Obstacles to Growth". Washington, D.C. World Bank.

__________ 2004b. "2003 Annual Review of Development Effectiveness: The Effectiveness of Bank Support for Policy Reform. Report 28290". Washington, D.C.: World Bank Operations Evaluation Department.

__________ 2005. "World Development Report 2005: A Better Investment Climate for Everyone". World Bank and Oxford University Press. Washington, D.C.

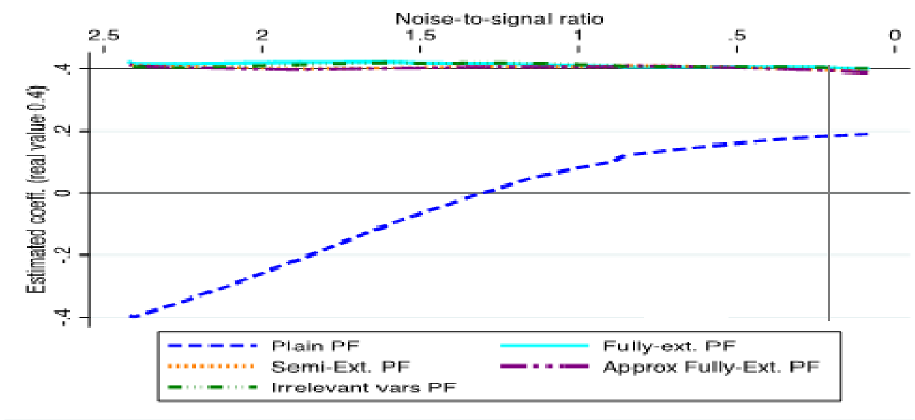
Appendix A: Monte Carlo Simulations Results.

FIGURE A.1: Results of simulation II, estimated coefficient of the OLS estimators of the effects of capital ic_1, ic_2 and ic_3 as a function of noise-to signal ratio (variance of the variable to variance of y)

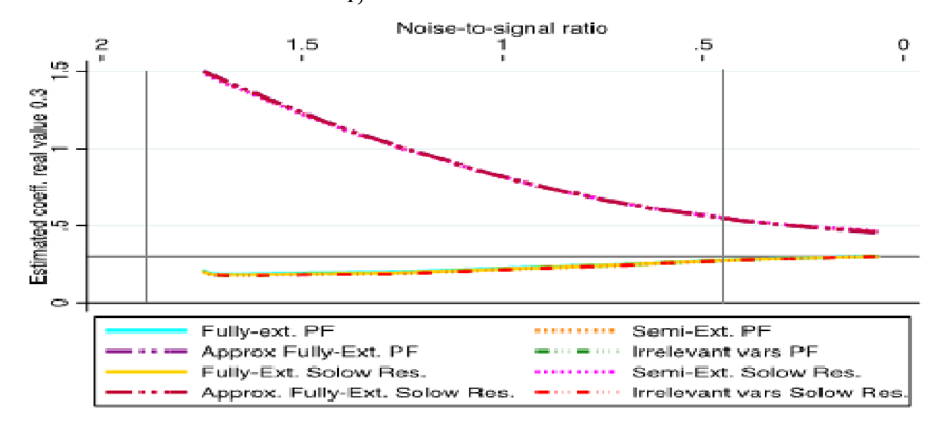
Lowess smooth fit between percentage bias and NtS ratio. Note: reverse scale in the noise-to-signal ratio axis

Capital empirical NtS ratio:0.24; Average empirical NtS of continuous IC variables:0.46, of dichotomous IC variables=1.94

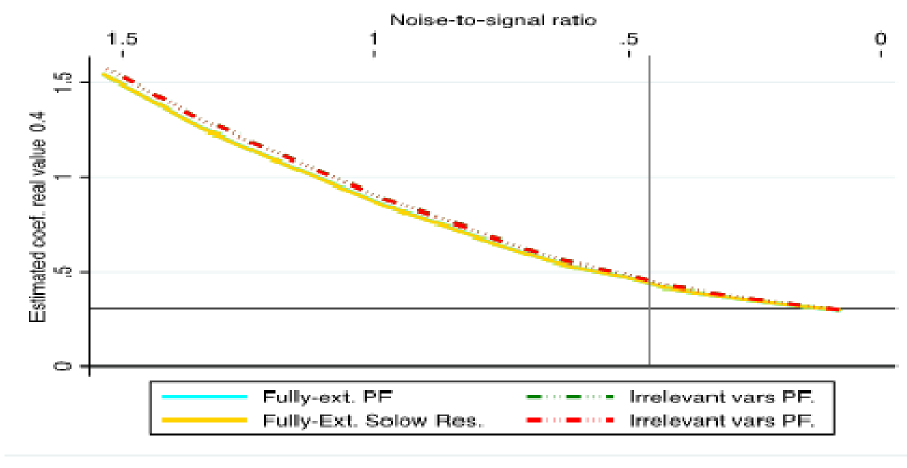
Capital, value in the DGP 0.4



ic_1, value in the DGP 0.3



ic_2, value in the DGP 0.3



ic_3, value in the DGP 0.09 and 0

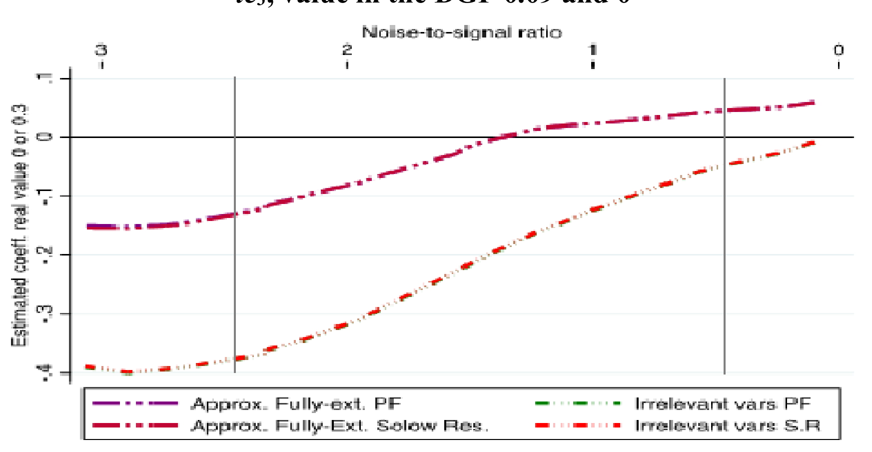
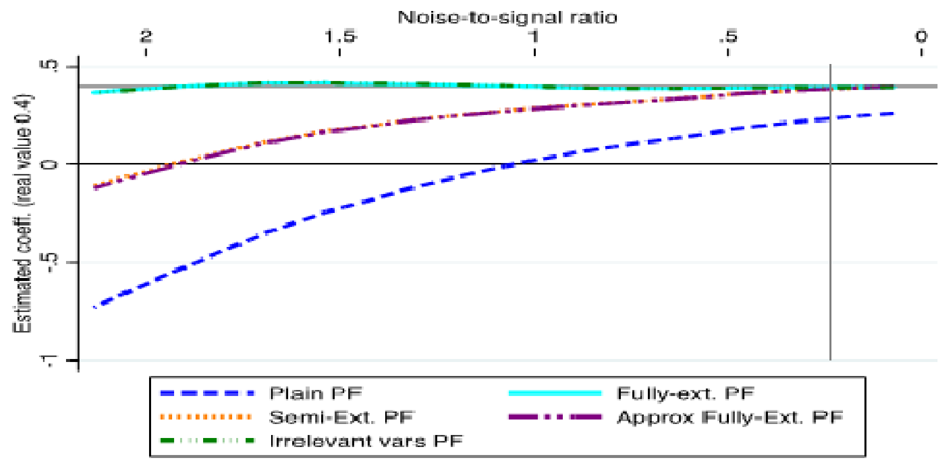


FIGURE A.2: Results of simulation IV, estimated coefficient of the OLS estimators of the effects of capital ic_1, ic_2 and ic_3 as a function of the noise-to signal ratio (variance of the variable to variance of y)

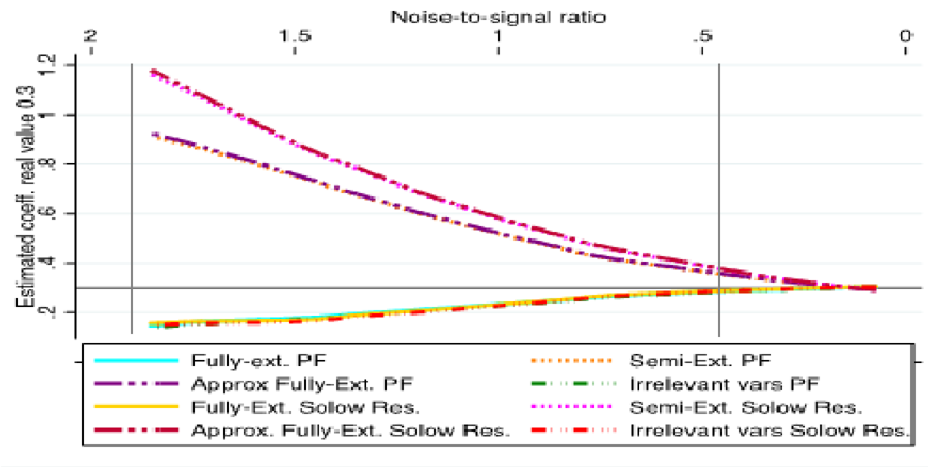
Lowess smooth fit between percentage bias and NtS ratio. Note: reverse scale in the noise-to-signal ratio axis

Capital empirical NtS ratio:0.24; Average empirical NtS of continuous IC variables:0.46, of dichotomical IC variables=1.94

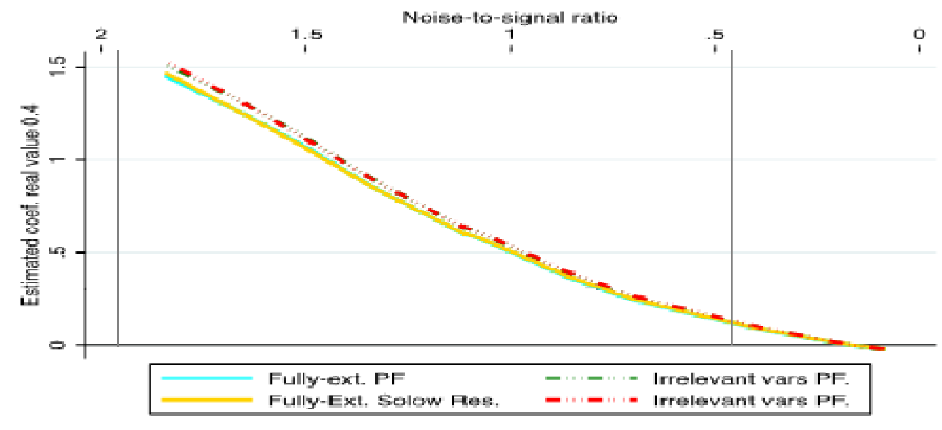
Capital, value in the DGP 0.4



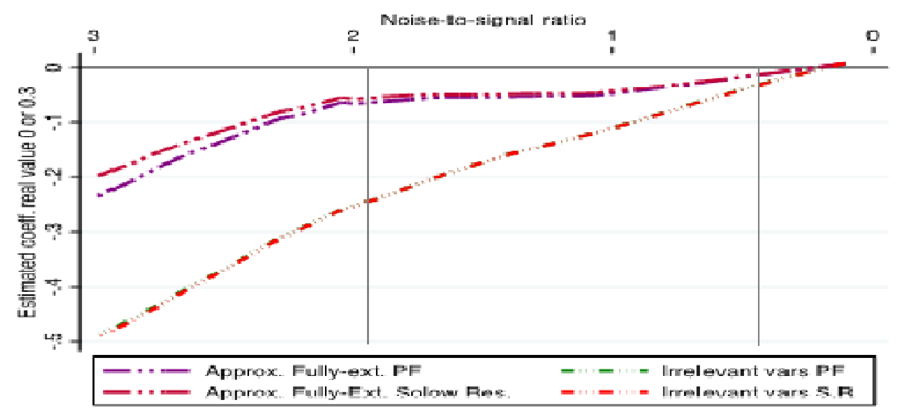
ic_1, value in the DGP 0.3



ic_2, value in the DGP 0



ic_3, value in the DGP 0.09 and 0



Appendix B: Figures from Costa Rica´s ICS

Figure B.1: Comparison of (log) TFP distributions and Kolmogorov-Smirnov tests of equality of distributions

Fig B.1a Including the constant term as part of the TFP

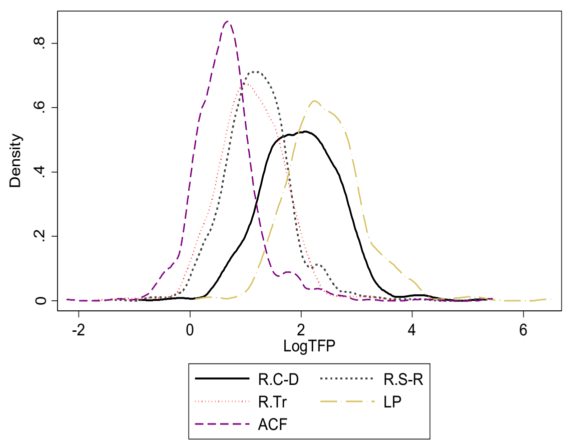
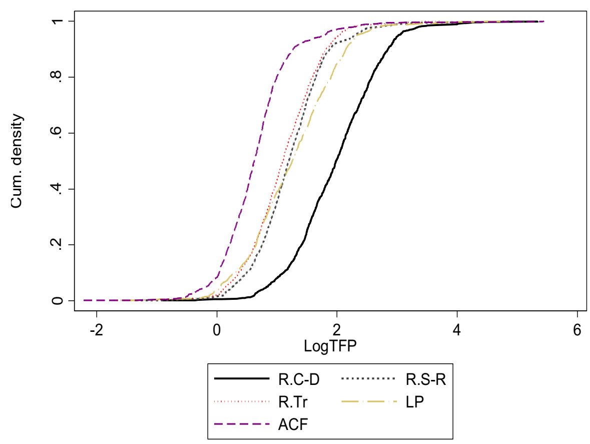
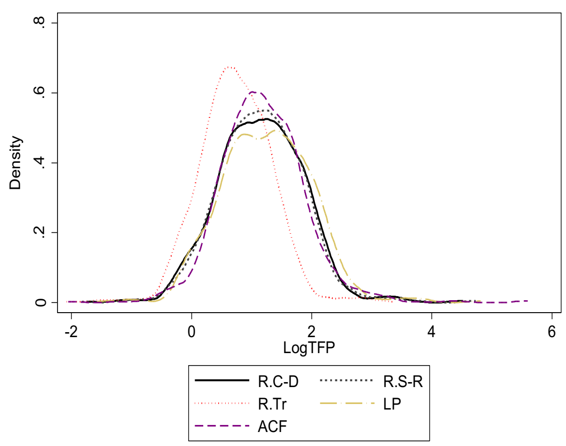
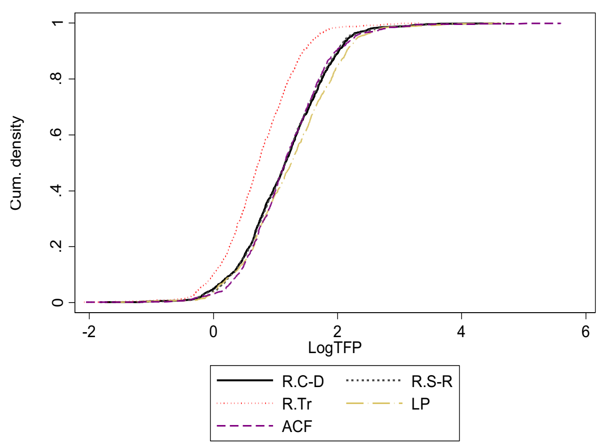
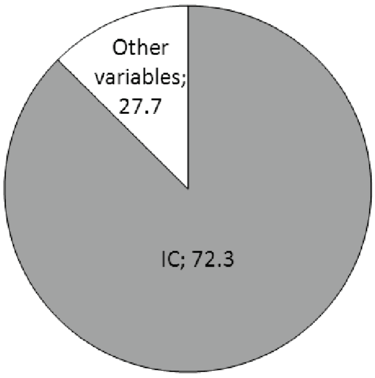
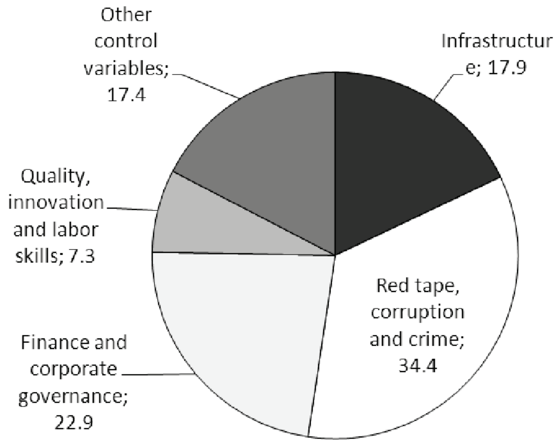
Kernel estimates of productivity densities	Cumulative densities of productivity distributions
	

Fig. B.1b Demeaned TFP: Excluding the constant term from TFP

Kernel estimates of productivity densities	Cumulative densities of productivity distributions
	

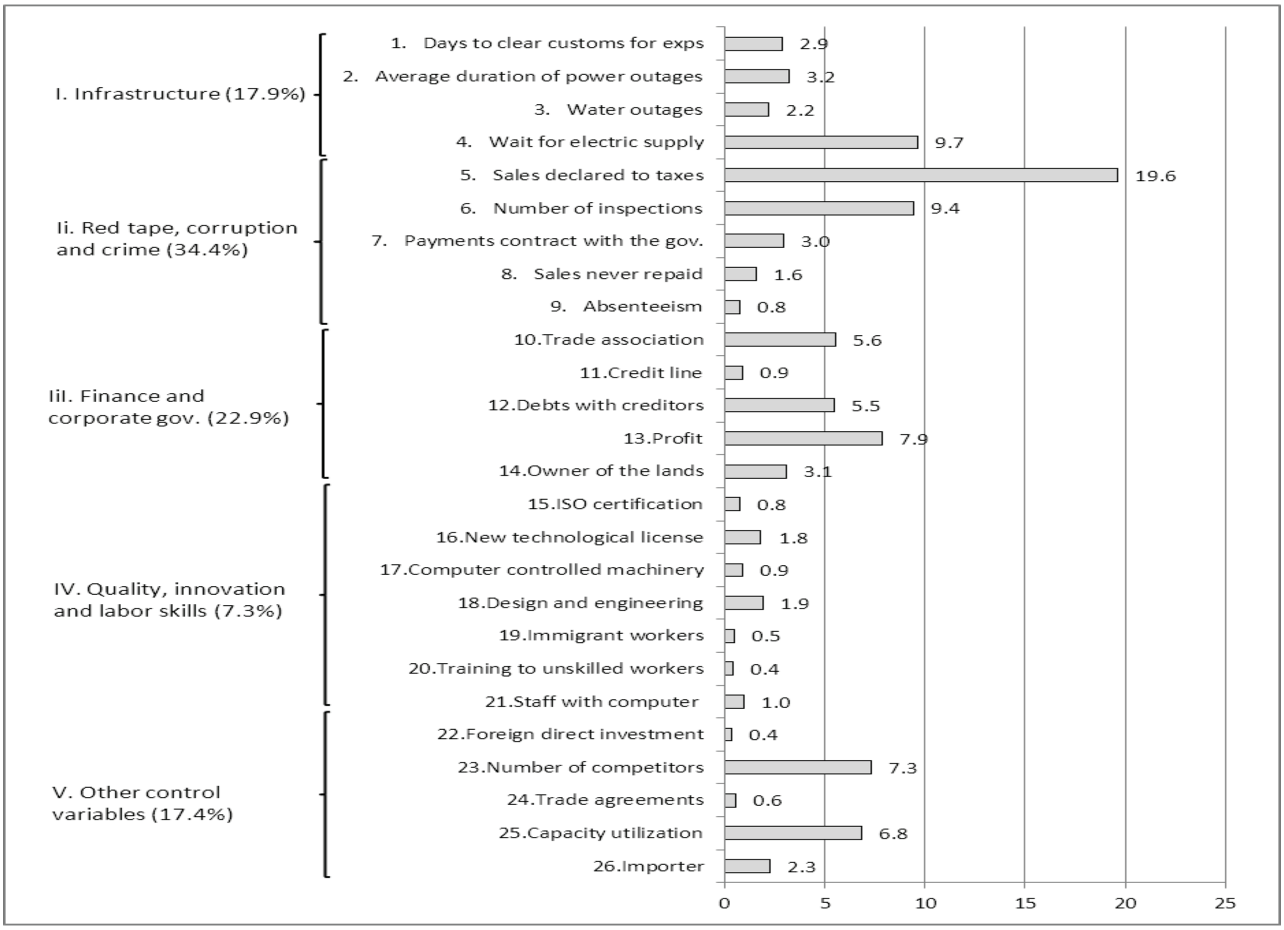
Note: Epanechnikov kernel. R.C-D=Restricted Cobb-Douglas, R.S-R=Restricted Solow residuals, R.Tr=Restricted Translog, LP=Levinson & Petrin and ACF=Ackerberg, Caves and Frazer.
Source: Authors' elaboration with IC data.

Figure B.2: Average percentage relative block contribution of the investment climate of Costa Rica to firms' TFP performance

Average percentage importance of the investment climate relative to other variables	Average percentage relative importance of the different investment climate blocks
	

Source: Authors' calculations with IC data of Costa Rica

Figure B.3: Average relative individual contribution of the investment climate variables of Costa Rica to firms' TFP performance



Source: Authors' calculations with IC data

Appendix C: Tables and Figures of Costa Rica

Table C.1: General information at plant level and Production Function variables

General Plant-Level Information of the Manufacturing Sector of Costa Rica	Industrial classification (8 sectors)	Food and beverages; textiles; apparels; wood and furniture; paper and edition; chemicals rubber and plastics; non-metallic products; machinery and equipment-metallic products.
	Regional classification	San José; Alajuela; Cartago; Heredia; Guanacaste; Puntarenas; Provincia Limón. Additional classification used in figures: Great Urban, Rest of Central Valley, Rest of the Country.
Production Function Variables	Sales	Used as the measure of output for the production function estimation. Sales are defined as total sales plus the changes in the inventories of finished goods. The series are deflated by using the Industrial Production Price Index, base 1999.
	Employment	Total number of permanent and temporal workers (full or part time).
	Total hours worked per year	Total number of employees multiplied by the average hours worked per year.
	Materials	Total costs of intermediate and raw materials used in production (excluding fuel). The series are deflated by using the Industrial Production Price Index, base 1999.
	Capital stock	Net book value of all fixed assets (log). The series are deflated by using the Industrial Production Price Index, base 1999.
	User cost of capital	The user cost of capital is defined in terms of the opportunity cost of using capital; it is defined as the long term interest rate in Costa Rica (more than 5 years) plus a depreciation rate of 20% minus the rate of growth of the consumption price index.
	Labor cost	Total expenditures on personnel, deflated by using the Industrial Production Price Index, base 1999.

Table C.2: List of Significant Investment climate (IC) variables

Name of the IC variables	**Definition of IC variables**
Group A of IC variables: Infrastructure	
Average duration of power outages Hours per day (logs) (AV)	Average duration of power outages suffered by the plant in hours during last fiscal year.
Average no. of days to clear customs for exports (logs) (AV)	Average number of days needed to clear customs for imports during last fiscal year.
Total number of water outages (logs) (AV)	Total number of water outages suffered by the plant during last fiscal year.
Average days waiting for an electricity supply (logs) (AV)	Number of days waiting for a public electric supply since the moment of the application to the day the service was received (number of days)
Group B of IC variables:. Red tape, Corruption and Crime	
Number of days spent in Inspection and Regulation related work Days (logs) (AV)	In the last year, total number of inspections regarding with taxes, employment, health control, municipal inspectors, etc.
Fraction of sales declared to IRS for tax purposes (Fraction of total sales) (AV)	Percentage of plant's total sales declared to taxes.
Dummy for payments to obtain a contract with the government (0 or 1) (AV)	Dummy that takes value 1 if firms in the main sector occasionally need to give gifts or make informal payments in order to get a contract with the government.
Percentage of sales never repaid (% of total sales) (AV)	Percentage of monthly total sales to private customers that were never repaid.
Number of days lost due to absenteeism (logs)	Days of production lost due to employees absenteeism during last year.
Group C of IC variables: Finance and Corporate Governance	
Dummy for firm belonging to a trade assoc. (0 or 1) (AV)	Dummy variable taking value 1 if the firm belongs to a trade association or trade chamber.
Dummy for credit line (0 or 1)	Dummy variable that takes value 1 if the plant reports that it has a credit line.
Dummy for debts with creditors (0 or 1) (AV)	Dummy variable that takes value 1 if the firm has any debt with suppliers.
Firm's profits after taxes as a percentage of total sales (% of total sales) (AV)	Firm's profits after taxes as a percentage of total sales.
Dummy for firm owning almost all the lands in which the plant operates (0 or 1)	Dummy variable that takes value 1 if the firm is the owner of almost all its lands.

* logs = logarithmic function; AV = Average value of the industry-region-size; (0, 1) = binary variable taking only 0 or 1 values.

Table C.2 (continued): List of Significant Investment climate (IC) variables

Group D of IC variables. Quality, innovation and labor skills	
Fraction computer-controlled machinery of total machinery (Fraction of total machinery)	Fraction computer-controlled machinery of total machinery (Fraction of total machinery)
Dummy for ISO quality certif. (0 or 1)	Dummy variable taking value 1 if the firm has any kind of ISO quality certification.
Dummy for new technological license (0 or 1) (AV)	Dummy variable that takes value 1 if the firm has acquired any new technology with important implications in the production process.
Number of plant's employees dealing with engineering and design (logs) (AV)	Number of plant's employees dealing with engineering and design (logs) (AV)
Percentage of immigrant workers (% of total staff) (AV)	Percentage of immigrant workers (% of total staff) (AV)
Percentage of unskilled workers receiving training (% of unskilled workers) (AV)	Percentage of unskilled workers receiving training (% of unskilled workers) (AV)
Percentage of staff using computer at job (% of total staff)	Percentage of staff using computer at job (% of total staff)
Group E of IC variables: Other control variables	
Dummy for foreign direct investment (0 or 1)	Dummy variable taking value 1 if any percentage of firm's share belongs to a foreign firm.
Number of competitors in plant's main market (logs) (AV)	Number of competitors in plant's main market (logs) (AV)
Dummy for benefit from free trade agreements with signed by the government (0 or 1)	Dummy for benefit from free trade agreements with signed by the government (0 or 1)
Percentage of capacity utilization (percentage)	Percentage of capacity used by the plant in average during last year.
Dummy for importer firm (0 or 1)	Dummy variable taking value one if the firm imports any share of its supplies.

* logs = logarithmic function; AV = Average value of the industry-region; (0, 1) = binary variable taking only 0 or 1 values.

Table C.3: OLS parameter estimation of extended restricted Cobb-Douglas production functions and Testing for the equal magnitude of the IC effects on eight alternative TFP measures

	Restricted Cobb-Douglas		Range of coeffs. based on the 8 specifications[1]	Statistically equal coeffs. over the 8 spec.[1]	Empirical Noise-to-Signal Ratios
	Coeff.	S.E	[min, max]	proportion	
Input-output elasticities					
Labor (logs)	0.312	(0.057)***			0.34
Materials (logs)	0.534	(0.046)***			0.23
Capital (logs)	0.121	(0.023)***			0.24
IC Coefficients (26 variables)					
1. Days to clear customs for exports (logs) (a)	-0.076	(0.03)**	[-0.090, -0.041]	7/8	0.72
2. Average duration of power outages (logs) (a)	-0.029	(0.009)***	[-0.035, -0.021]	7/8	0.13
3. Water outages (logs) (a)	-0.217	(0.097)**	[-0.302, -0.164]	7/8	1.67
4. Wait for electric supply (logs) (a)	-0.128	(0.026)***	[-0.144, -0.059]	7/8	0.31
5. Sales declared to taxes (%) (a)	0.01	(0.003)***	[0.003, 0.012]	5/8	0.04
6. Number of inspections (logs) (a)	-0.326	(0.075)***	[-0.36, -0.262]	7/8	1.05
7. Payments to obtain a contract with the government (0 or 1) (a)	0.394	(0.198)*	[0.178, 0.464]	7/8	3.20
8. Sales never repaid (%) (a)	-0.016	(0.004)***	[-0.016, -0.002]	5/8	0.12
9. Absenteeism (logs)	-0.042	(0.019)**	[-0.048, -0.0213]	8/8	0.45
10. Trade association (0 or 1) (a)	0.447	(0.141)***	[0.284, 0.568]	8/8	2.42
11. Credit line (0 or 1)	0.07	-0.048	[0.022, 0.088]	7/8	1.07
12. Debts with creditors (0 or 1) (a)	0.276	-0.203	[0.027, 0.395]	7/8	2.92
13. Profit (%) (a)	0.018	(0.007)***	[0.008, 0.021]	7/8	0.07
14. Owner of the lands (0 or 1)	-0.158	(0.055)***	[-0.171, -0.152]	7/8	1.13
15. ISO certification (0 or 1)	0.301	(0.089)***	[0.18, 0.334]	7/8	1.83
16. New technological license (0 or 1) (a)	0.196	-0.141	[0.083, 0.287]	8/8	2.48
17. Computer controlled machinery (%)	0.003	(0.001)*	[0.002, 0.003]	7/8	0.02
18. Design and engineering (logs) (a)	0.031	(0.014)**	[0.017, 0.044]	8/8	0.29
19. Immigrant workers (%) (a)	-0.133	(0.063)**	[-0.198, -0.023]	6/8	1.66
20. Training to unskilled workers (%) (a)	0.004	-0.004	[0.002, 0.011]	8/8	0.10
21. Staff with computer (%)	0.002	-0.002	[0.001, 0.003]	8/8	0.02
22. Foreign direct investment (0 or 1)	0.156	-0.121	[0.117, 0.183]	7/8	1.85
23. Number of competitors (logs) (a)	0.125	(0.033)***	[0.107, 0.164]	8/8	0.71
24. Trade agreements (0 or 1)	0.109	-0.078	[0.026, 0.132]	6/8	1.36
25. Capacity utilization (%)	0.003	(0.001)***	[0.002, 0.003]	7/8	0.03
26. Importer (0 or 1)	0.22	(0.077)***	[0.088, 0.259]	6/8	1.11
Number of observations	985				

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%.
Robust standard errors in parentheses, also corrected for correlation within clusters defined by industry and region of the establishment. Each of the 8 specifications includes a constant term, a set of industry and year dummies with parameters not reported here.
[1]The 8 specifications considered are: 1) Restricted Cobb-Douglas, 2) Unrestricted Cobb-Douglas, 3) Restricted Translog, 4) Unrestricted Translog, 5) Restricted Solow Residual, 6) Unrestricted Solow residual, 7) Levinsohn and Petrin, 8) Ackerberg, Caves and Frazer.
(a) IC variables expressed as industry-region-size average. Other IC variables information is also indicated in parentheses; logs, % or dummy (0 or 1).
Weak Instruments test: The F statistic associated to the null hypothesis that the coefficients of the 26 IC variables are all jointly equal to zero is F=19.54, with a p-value approximately equal to zero.
Source: Authors' calculations with IC data.

Table C.4: Testing the Equal Magnitude of the coefficients of 26 IC variables across 8 Alternative TFP measures

IC Coefficients (26 variables)	[1] Restricted Cobb-Douglas	[2] Unrestricted Cobb-Douglas		[3] Restricted Solow Residual		[4] Unrestricted Solow Residual		[5] Restricted Translog		[6] Unrestricted Translog		[7] Levinson & Petrin		[8] AC&F		[9] C-D + Polynomial	
	Coeff	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
1. Days to clear customs for exps (a)	-0.076	-0.076	0.99	-0.079	0.652	-0.090	0.117	-0.064	0.417	-0.041	0.185	-0.043**	0.01	-0.084	0.662	-0.019**	0.027
2. Average duration of power outages (a)	-0.029	-0.030	0.834	-0.029	0.734	-0.027	0.479	-0.022	0.234	-0.021	0.485	-0.035	0.505	-0.023*	0.085	-0.027	0.764
3. Water outages (a)	-0.217	-0.207	0.856	-0.202	0.566	-0.237	0.588	-0.164	0.361	-0.215	0.977	-0.243	0.84	-0.302**	0.042	-0.078**	0.026
4. Wait for electric supply (a)	-0.128	-0.106	0.423	-0.127	0.736	-0.118	0.26	-0.144	0.128	-0.059**	0.027	-0.133	0.896	-0.122	0.61	-0.081*	0.057
5. Sales declared to taxes (a)	0.01	0.010	0.697	0.010	0.546	0.010	0.933	0.005*	0.058	0.006*	0.097	0.009**	0.042	0.012**	0.016	0.003***	0.003
6. Number of inspections (a)	-0.326	-0.326	0.992	-0.317	0.346	-0.345*	0.096	-0.262	0.164	-0.281	0.548	-0.336	0.568	-0.36	0.242	-0.170***	0.006
7. Payments contract government (a)	0.394	0.240	0.218	0.388	0.85	0.418	0.58	0.276	0.342	-0.102***	0.007	0.464	0.919	0.342	0.39	0.178	0.148
8. Sales never repaid (a)	-0.016	-0.014	0.603	-0.018	0.342	-0.015	0.699	-0.007**	0.021	-0.002***	0.007	-0.014	0.527	-0.011***	0.006	-0.004***	0.009
9. Absenteeism	-0.042	-0.023	0.117	-0.042	0.945	-0.046	0.497	-0.036	0.276	-0.021	0.166	-0.041	0.817	-0.042	0.972	-0.048	0.657
10. Trade association (a)	0.447	0.568	0.155	0.460	0.588	0.403	0.148	0.284	0.163	0.453	0.96	0.315	0.475	0.442	0.935	0.284	0.144
11. Credit line	0.07	0.039	0.38	0.088	0.302	0.054	0.478	0.037	0.309	0.033	0.4	0.046	0.689	0.022	0.053*	0.061	0.839
12. Debts with creditors (a)	0.276	0.330	0.602	0.249	0.539	0.276	0.996	0.107	0.301	0.156	0.617	0.283***	0.004	0.395	0.188	0.027	0.113
13. Profit (a)	0.018	0.014	0.316	0.017	0.306	0.018	0.869	0.015	0.327	0.008	0.125	0.021*	0.059	0.019	0.624	0.008*	0.06
14. Owner of the lands	-0.158	-0.158	0.995	-0.152	0.684	-0.154	0.826	-0.156	0.947	-0.161	0.954	-0.171***	0.000	-0.16	0.933	-0.159	0.982
15. ISO certification	0.301	0.180	0.054	0.334	0.289	0.283	0.586	0.264	0.393	0.207	0.153	0.247	0.2	0.183**	0.011	0.218	0.107
16. New technological license (a)	0.196	0.116	0.376	0.216	0.524	0.211	0.684	0.287	0.221	0.215	0.874	0.179	0.107	0.083	0.102	0.205	0.923
17. Computer controlled machinery	0.003	0.003	0.775	0.003	0.438	0.003	0.359	0.003	0.699	0.001	0.291	0.002***	0.001	0.002	0.197	0.002	0.482
18. Design and engineering (a)	0.031	0.017	0.134	0.029	0.451	0.033	0.622	0.038	0.392	0.003**	0.03	0.044	0.461	0.025	0.447	0.019	0.207
19. Immigrant workers (a)	-0.133	-0.023**	0.022	-0.124	0.529	-0.173**	0.035	-0.125	0.736	-0.035	0.109	-0.198**	0.021	-0.147	0.548	-0.083	0.328
20. Training to unskilled workers (a)	0.004	0.007	0.301	0.004	0.939	0.004	0.91	0.003	0.501	0.011	0.021	0.002	0.898	0.007	0.161	0.006	0.429
21. Staff with computer	0.002	0.002	0.412	0.002	0.748	0.002	0.134	0.002	0.811	0.003	0.328	0.002	0.42	0.001	0.527	0.002	0.511
22. Foreign direct investment	0.156	0.183	0.644	0.173	0.546	0.143	0.713	0.119	0.601	0.073	0.36	0.138**	0.018	0.117	0.465	0.127	0.736
23. Number of competitors (a)	0.125	0.107	0.478	0.119	0.424	0.137	0.238	0.164	0.171	0.100	0.483	0.161	0.992	0.135	0.61	0.134	0.75
24. Trade agreements	0.109	0.095	0.746	0.132	0.301	0.057	0.114	0.078	0.478	0.112	0.955	0.094***	0.000	0.026**	0.032	0.110	0.973
25. Capacity utilization	0.003	0.003	0.735	0.003	0.478	0.003	0.613	0.003	0.766	0.003	0.556	0.003*	0.065	0.002	0.159	0.003	0.845
26. Importer	0.22	0.190	0.357	0.259	0.235	0.187	0.375	0.242	0.538	0.185	0.422	0.189***	0.000	0.088***	0.000	0.223	0.943

Note: the equality of the IC coefficients is performed relative to the Restricted Cobb-Douglas (RCD) specification, column [1]. Therefore, the test statistics and p-values measure whether the estimated IC coefficients are statistically different from those of the RCD. The 8 specifications considered are: 1) Restricted Cobb-Douglas, 2) Unrestricted Cobb-Douglas, 3) Restricted Translog, 4) Unrestricted Translog, 5) Restricted Solow Residual, 6) Unrestricted Solow residual, 7) Levinsohn and Petrin, 8) Ackerberg, Caves and Frazer.* means the corresponding estimated coefficient is significantly different at the 10% level of confidence, ** if it is different at 5% and *** at 1%.
Source: Authors' calculations with IC data.

Table C.5: Summary of equality of coefficients tests across 8 TFP specifications and for each of 26 IC explanatory variables

Percentages and number of IC coincidences with respect to the baseline model (Restricted Cobb-Douglas)

	Percentage of sign coincidences	Number of statistically equal IC coeffs. (% in parentheses)
A. Unrestricted Cobb-Douglas	100%	25 (96.1%)
B. Restricted Solow residual	100%	26 (100%)
C. Unrestricted Solow Residual	100%	24 (92.3%)
D. Restricted Translog	100%	24 (92.3%)
E. Unrestricted Translog	87.5%	21 (80.1%)
F. Levinson &Petrin (L&P)	100%	15 (57.7%)
G. AC&F	100%	18 (53.9%)
H. Cobb-Douglas + AC&F Polynomial	100%	-

Notes: [1]The null hypothesis is all the 26 IC coefficients of the alternative TFP specifications are equal to those of the restricted Cobb-Douglas. Under the null the test statistic is distributed as $F_{26,47.}$

Source: Authors' calculations with IC data.

Table C.6: Comparison of productivity (TFP) distributions: mean, standard deviations and Kolmogorov-Smirnov[1] (K-S) Tests of equality of distributions

	TFP with constant term				TFP without constant term (Demeaned)			
TFP measure	**Mean**	**S.D**	**Corr.**	**K-S test (p-value)**	**Mean**	**S.D**	**Corr.**	**K-S test (p-value)**
1. Restricted Cobb-Douglas	1.985	0.714	-	-	1.156	0.714	-	-
2. Unrestricted Cobb-Douglas	1.487	0.644	0.951	0.478 (0.000)	1.084	0.644	0.949	0.018 (0.997)
3. Restricted Solow residual	1.222	0.632	0.865	0.501 (0.000)	1.154	0.704	0.945	0.053 (0.128)
4. Unrestricted Solow Residual	1.196	0.647	0.869	0.31 (0.000)	1.098	0.714	0.944	**0.092 (0.000)**
5. Restricted Translog	1.105	0.601	0.932	0.51 (0.000)	0.734	0.601	0.952	**0.276 (0.000)**
6. Unrestricted Translog	-0.897	0.535	0.853	0.982 (0.000)	0.742	0.535	0.873	**0.308 (0.000)**
7. Levinson & Petrin	2.44	0.702	0.835	0.933 (0.000)	1.250	0.702	0.985	0.07 (0.021)
8. AC&F	0.657	0.626	0.82	0.716 (0.000)	1.187	0.626	0.855	0.048 (0.212)

Notes: The Kolmogorov-Smirnov (K-S) test compares, pairwise, all the alternative TFP distributions with the one of the Restricted Cobb-Douglas. Null hypothesis: Each alternative TFP measure is generated by the same distribution as the Restricted Cobb-Douglas TFP case.

Source: Authors' elaboration with IC data.